Some Topics in Large Deviations Theory for Stochastic Dynamical Systems

Jiang Chen

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Statistics and Operations Research.

Chapel Hill
2013

Approved by:

Amarjit Budhiraja

Shankar Bhamidi

Jan Hannig

Chuanshu Ji

Vladas Pipiras

Abstract

JIANG CHEN: Some Topics in Large Deviations Theory for Stochastic Dynamical Systems
(Under the direction of Amarjit Budhiraja)

In this dissertation, we study *large deviations* problems for stochastic dynamical systems. First, we consider a family of Stochastic Partial Differential Equations (SPDE) driven by a Poisson Random Measure (PRM) that are motivated by problems of chemical/pollutant dispersal. We established a Large Deviation Principle (LDP) for the long time profile of the chemical concentration using techniques based on variational representations for nonnegative functionals of general PRM. Second, we develop a LDP for small Poisson noise perturbations of a general class of deterministic infinite dimensional models. SPDEs driven by PRM have been proposed as models for many different physical systems. The approach taken here, which is based on variational representations, reduces the proof of the LDP to establishing basic qualitative properties for controlled analogues of the underlying stochastic system. Third, we study stochastic systems with two time scales. Such multiscale systems arise in many applications in engineering, operations research and biological and physical sciences. The models considered in this dissertation are usually referred to as systems with "full dependence", which refers to the feature that the coefficients of both the slow and the fast processes depend on both variables. We establish a LDP for such systems with degenerate diffusion coefficients.

The last part of this dissertation focuses on numerical schemes for computing invariant measures of reflected diffusions. Reflected diffusions in polyhedral domains are commonly used as approximate models for stochastic processing networks in heavy traffic. Stationary distributions of such models give useful information on the steady state performance of the corresponding stochastic networks and thus it is important to develop reliable and efficient algorithms for numerical computation of such distributions. We propose and analyze a Monte-Carlo scheme based on an Euler type discretization. We prove an almost sure consistency of the appropriately weighted empirical measures constructed from the simulated discretized reflected diffusion to the true diffusion model. Rates of convergence are also obtained for certain class of test functions. Some numerical examples are presented to illustrate the applicability of this approach.

Acknowledgements

This is a great opportunity to express my gratitude to a number of individuals who have helped me in achieving the task of completing this dissertation.

My deepest gratitude is to my advisor, Dr. Amarjit Budhiraja. I am deeply indebted to him for his supervision, encouragement, patience, and rigorous yet pleasant training during my Ph.D. study. It was Dr. Budhiraja who brought me into the probability world, who built my background in probability through rigorous training and countless hours of face-to-face instructions. I am also grateful to him for his great patience in checking over all the technical details and grammars in countless revisions of the work. Finally, I am thankful to him for his financial support, with which I can fully concentrate on my study and research.

Also, I would like to acknowledge my committee members. In particular, I am grateful to Dr. Shankar Bhamidi and Dr. Vladas Pipiras for their valuable suggestions to my future work. I thank Dr. Jan Hannig, for teaching me the fundamentals and for his wise advice. I am also indebted to Dr. Chuanshu Ji for his kind help and encouragement.

I would like to acknowledge Dr. Ross Leadbetter, Dr. Andrew Nobel, Dr. Haipeng Shen and Dr. Steve Marron for numerous discussions and lectures on related topics that helped me improve my knowledge in the area.

I am also thankful to the following staff at University of North Carolina-Chapel Hill, for their various forms of support during my graduate study–Ms. Alison Kieber, Ms. Elizabeth Moffitt-Johnson, and Ms. Charlotte Rogers.

I am grateful to many friends. Their support and care helped me overcome setbacks and stay focused on my graduate study. I greatly value their friendship and I deeply appreciate their belief in me. My gratitude especially goes to Mr. Xuan Wang, for numerous inspirational discussions, group studies and support.

Most importantly, none of this would have been possible without the love and patience of my family. My family has been a constant source of love, concern, support and strength all these years. I would like to express my heart-felt gratitude to my family.

Table of Contents

List of Notation and Symbols

$\mathbb{N}$	Set of natural numbers
$\mathbb{N}_0$	Set of non-negative integers
$\mathbb{R}$	Set of real numbers
$\mathbb{R}_+$	Set of positive real numbers
$\mathbb{R}^d$	Set of d-dimensional real vectors
$\mathbb{Z}$	Set of integers
$\mathcal{B}(\mathcal{E})$	Borel σ-field on a topological space $\mathcal{E}$
$C_b(\mathbb{X})$	The collection of bounded continuous functions from space $\mathbb{X}$ to $\mathbb{R}$
$M_b(\mathbb{X})$	The collection of bounded $\mathcal{B}(\mathbb{X})/\mathcal{B}(\mathbb{R})$ measurable maps
$\mathcal{P}(\mathbb{X})$	Collection of all probability measures on space $\mathbb{X}$
$\mathcal{M}_{FC}(\mathbb{X})$	Collection of all measures which are finite on compacts on space $\mathbb{X}$
$\mathcal{M}_F(\mathbb{X})$	Collection of all finite measures on space $\mathbb{X}$ endowed with the usual topology of weak convergence
$\mathbb{M}(\mathbb{L})$	The space of real functions on $\mathbb{L}$, for a finite set $\mathbb{L}$
$C([0,T]:\mathbb{X})$	The space of continuous functions from $[0,T]$ to $\mathbb{X}$
$D([0,T]:\mathbb{X})$	The space of right continuous functions with left limits from $[0,T]$ to $\mathbb{X}$
$\|\|f\|\|_\infty$	For $f \in C([0,T]:\mathbb{X})$, $\|\|f\|\|_\infty = \sup_{x\in[0,T]} \|f(x)\|$ For a bounded $\mathbb{R}^d$ valued function f on S, $\|\|f\|\|_\infty = \sup_{x\in S} \|f(x)\|$
$X_n \Rightarrow X$	Weak convergence of stochastic process X_n to X
$\mu_n \Rightarrow \mu$	Weak convergence of probability measures μ_n to μ
$L^2(\mathcal{E},\nu;H)$	The space of measurable functions f from $\mathcal{E}$ to H such that $\int_{\mathcal{E}} \|\|f(v)\|\|^2\nu(dv) < \infty$, where $\|\|\cdot\|\|$ is the norm on H, for a measure ν on $\mathcal{E}$ and a Hilbert space H
x_t and $x(t)$	The evaluation of x at $t \in [0,T]$ for a function $x:[0,T]\to\mathcal{E}$, and similar convention will be followed for stochastic processes

$\{X^\epsilon\}$ is tight	The distributions of X^ϵ are tight
ODE	Ordinary differential equation
PDE	Partial differential equation
SPDE	Stochastic partial differential equation
PRM	Poisson random measure
LDP	Large deviation principle

Chapter 1

Introduction

This dissertation contains two distinct parts. The first part, which is the main part (including Chapter 2-5), concerns large deviation theory for stochastic dynamical systems, while the second discusses a problem in stochastic numerics.

The theory of large deviations is concerned with the study of asymptotic exponential decay rate of probabilities of "rare events". In a typical setting, one is given a sequence of random variables $\{X_n\}$ with values in some Polish space $(\mathcal{E}, d)$ which, as $n \to \infty$, converges to a non-random limit $x \in \mathcal{E}$. The main problem of interest is to obtain precise upper and lower asymptotic bounds (as $n \to \infty$) on probabilities of deviations of X_n from its limit value x, e.g. on quantities such as $\mathbb{P}(d(X_n, x) \geq \alpha)$. A systematic treatment of this asymptotic study is given by establishing a *Large Deviation Principle* (LDP) which gives precise exponential decay rates for probabilities of the above form in terms of a suitable rate function.

The first part of this dissertation studies large deviation results for certain families of stochastic dynamical systems with jumps. Many models in probability and stochastic dynamics are given in terms of noise processes that are described in terms of Poisson random measures (PRM) and/or Brownian motions (BM). A promising approach based on certain variational formulas to treat large deviation problems for such stochastic systems has been initiated in [18]. These variational formulas are the starting point of my work. We collect these formulas together with other background results in Chapter 2.

We study two sets of applications of the variational representations of [15] to large deviation problems for stochastic systems. The first is to the study of small noise stochastic partial differential equations (SPDE) with Poisson noise. This is contained in Chapters 3 and 4 of the dissertation. A family of SPDE motivated by problems of chemical/pollutant dispersal is discussed in Chapter 3. In Chapter 4, we studied a rather general family of SPDE models driven by PRM, and established a large deviation result. The second application, studied in Chapter 5, concerns stochastic averaging problems for two time scale stochastic differential equations (SDE) with full dependence.

Although now there are many papers that treat large deviation problems for SPDEs driven by Gaussian noises (see [15] and references therein), there are almost no results available that systematically cover the setting of SPDEs with jumps. SPDEs driven by PRM have been proposed as models for many different physical systems, where they are viewed as a refinement of a corresponding noiseless partial differential equation (PDE). For example, they have been used to develop models for neuronal activity that account for synaptic impulses occurring randomly, both in time and at different locations of a spatially extended neuron. Other applications arise in chemical reaction-diffusion systems and stochastic turbulence models. We are interested in the study of probabilities of deviations of the stochastic PDE from the associated deterministic PDE. A systematic framework for such a study is through the theory of large deviations. This is the topic of Chapters 3 and 4 of the dissertation.

In Chapter 3, we consider a family of SPDE driven by a Poisson random measure that is motivated by problems of chemical/pollutant dispersal. These equations (taken from [52]) are stochastic versions of well studied convection-diffusion equations from hydrology literature for spread of a contaminant in a reservoir. We are interested in the long time profile of the contaminant concentration. In particular, we study the probability of deviations from the nominal behavior determined by the law of large numbers. The model is treated separately in two settings according to values of a parameter in Section 3.3 and Section 3.4 respectively. In both cases, estimates on probabilities of deviations are obtained by establishing a suitable large deviation principle.

In Chapter 4, guided by the particular problem in Chapter 3, we develop the large deviation theory for small Poisson noise perturbations of a general class of deterministic infinite dimensional models. In typical settings, the solutions of these stochastic evolution equations are not smooth. In fact in many applications of interest they are not even random fields (that is, function valued), and therefore an appropriate framework is given through the theory of generalized functions. In this chapter, we extend the approach based on variational representations [18] that has been very successful for obtaining large deviation results for infinite dimensional systems driven by Brownian noises to SPDE models driven by PRMs. As in the Brownian case the focus here is on the small noise problem, which in the Poisson setting means that the jump intensity is $O(\epsilon^{-1})$ and jump sizes are $O(\epsilon)$, where ϵ is a small parameter. We consider a rather general family of models of the form

$$X_t^\epsilon = X_0^\epsilon + \int_0^t A(s, X_s^\epsilon)ds + \epsilon \int_0^t \int_{\mathbb{X}} G(s, X_{s-}^\epsilon, v)\tilde{N}^{\epsilon^{-1}}(dsdv), \tag{1.0.1}$$

where $\tilde{N}^{\epsilon^{-1}}$ is the compensated PRM, associated with a PRM $N^{\epsilon^{-1}}$ on $[0,T]\times\mathbb{X}$ with a σ-finite mean measure $\epsilon^{-1}\lambda_T \otimes \nu$, where λ_T is the Lebesgue measure on $[0,T]$ and ν is a σ-finite measure on the locally compact space $\mathbb{X}$.

As noted previously, a key issue with a Poisson noise model is the selection of an appropriate state space, since there is little spatial regularity. However, many of these foundational issues have been satisfactorily resolved in [52], where pathwise existence and uniqueness of SPDE of the form (1.0.1) are treated under rather general conditions. We find that the estimates developed in [52] for establishing the well-posedness of the equation are precisely the ones that are key for the proof of the large deviation result as well. We provide verifiable conditions on the model data in (1.0.1) under which a large deviations principle holds. The paper based on this chapter (joint work with A. Budhiraja and P. Dupuis) has appeared in *Stochastic Processes and their Applications* [11].

Chapter 5 is devoted to stochastic systems with two time scales. In this chapter, we study large deviation properties of stochastic differential equations with coefficients that are governed by rapidly oscillating pure jump processes. Such systems arise in a variety of applications from stochastic networks, mathematical finance, stochastic control and optimization; see for example [88] and references therein. The precise mathematical model consists of a two component Markov process (X^ϵ, Y^ϵ) with values in $\mathbb{G} = \mathbb{R}^d \times \mathbb{L}$, where $\mathbb{L}$ is a finite set which without loss of generality we take to be a finite additive group, and $\epsilon > 0$ is a small parameter. The process X^ϵ between consecutive jumps of Y^ϵ is a diffusion with coefficients $b(x, y)$ and $a(x, y)$, and Y^ϵ is a pure jump process described in terms of the jump intensity function $c(x, y)$ and the transition probability kernel $R(x, y, d\bar{y})$. More precisely, if $(X^\epsilon(0), Y^\epsilon(0)) = (x, y) \in \mathbb{R}^d \times \mathbb{L}$, denoting by τ the first jump instant of Y^ϵ, on $[0, \tau)$, the process X^ϵ satisfies

$$dX^\epsilon(t) = b(X^\epsilon(t), y)dt + \sqrt{\epsilon}a(X^\epsilon(t), y)dW(t),$$

where W is a d dimensional Brownian motion. Furthermore,

$$\mathbf{P}(\tau > t \mid \sigma\{X^\epsilon(s), s \leq t\}) = \exp\left\{-\frac{1}{\epsilon}\int_0^t c(X^\epsilon(s), y)ds\right\}$$

and

$$\mathbf{P}\left(Y^\epsilon(\tau) \in d\bar{y} \mid X^\epsilon(\tau-) = x, Y^\epsilon(\tau-) = y\right) = R(x, y, d\bar{y}).$$

Thus the pair (X^ϵ, Y^ϵ) describes a jump-diffusion, where the diffusion component (i.e. X^ϵ) has "small noise" while the jump component Y^ϵ has jumps at rate $O(1/\epsilon)$.

It is well known (cf. [73]) that as $\epsilon \to 0$, X^ϵ converges to the solution of an ODE $\dot{x} = \hat{b}(x)$, where for fixed x, $\hat{b}(x)$ is given in terms of the invariant distribution of a $\mathbb{L}$ valued Markov process whose jump intensity function and transition kernel are $c(x, \cdot)$ and $R(x, \cdot, \cdot)$ respectively. In this chapter we establish a LDP for $\{X^\epsilon\}_{\epsilon>0}$ as $\epsilon \to 0$. One of the key challenges, particularly in the proof of the lower bound, is to handle the possible degeneracy of the diffusion coefficient.

Starting point of our analysis is a representation for the pair (X^ϵ, Y^ϵ) through a stochastic evolution equation given in terms of a suitable Poisson random measure.

$$\begin{aligned} dX^\epsilon(t) &= b(X^\epsilon(t), Y^\epsilon(t))dt + \sqrt{\epsilon} a(X^\epsilon(t), Y^\epsilon(t))dW(t),\ X^\epsilon(0) = x_0; \\ dY^\epsilon(t) &= \int_{r\in[0,1]} k(X^\epsilon(t), Y^\epsilon(t), r) N^{1/\epsilon}(dr \times dt),\ Y^\epsilon(0) = y_0. \end{aligned}$$

The function k and the intensity of $N^{1/\epsilon}$ are of course related to the functions $c(\cdot, \cdot)$ and $R(\cdot, \cdot, \cdot)$. This representation enables us to use the variational formulas for functionals of Brownian motion and PRM obtained in [18]. Using techniques from the theory of weak convergence and stochastic averaging, we then establish a LDP for $\{X^\epsilon\}_{\epsilon>0}$.

The second part of this dissertation consists of Chapter 6, in which we study a numerical scheme for invariant distributions of constrained diffusions. Constrained diffusion processes in polyhedral domains have been proposed as approximate models for critically loaded stochastic processing networks. Many performance measures for stochastic networks are formulated to capture the long term behavior of the system and a key object involved in the computation of such measures is the corresponding steady state distribution. There are now several results that prove, for certain generalized Jackson network models, the convergence of steady state distributions of stochastic networks to those of the associated limit diffusions. Indeed, one of the main motivations for introducing diffusion approximations in the study of stochastic processing systems is the expectation that diffusion models are easier to analyze than their stochastic network counterparts. Then an important question is how to compute the stationary distribution of reflected diffusions. Classical results of Harrison and Williams [48] show that under certain geometric conditions on the underlying problem data, stationary densities of reflected Brownian motions have explicit product form expressions. However, once one moves away from this special family of models there are no explicit formulas and thus one needs to use numerical procedures.

The objective of this chapter is to propose and study the performance of one such numerical procedure for computing stationary distributions of reflected diffusions in polyhedral domains. We propose and analyze a Monte-Carlo scheme based on an Euler type discretization of the reflected SDE using a single sequence of time discretization steps which decrease to zero as time approaches infinity. Appropriately weighted empirical measures constructed from the simulated discretized reflected diffusion are proposed as approximations for the invariant probability measure of the true diffusion model. Almost sure consistency results are established that in particular show that weighted averages of polynomially growing continuous functionals evaluated on the discretized simulated system converge a.s. to the corresponding integrals with respect to the invariant measure. Proofs rely on constructing suitable Lyapunov functions for tightness and uniform integrability and characterizing almost sure limit points through an extension of Echeverria's criteria for reflected diffusions. Regularity properties of the underlying Skorohod problems play a key role in the proofs. Rates of convergence for suitable families of test functions are also obtained. A key advantage of Monte-Carlo methods is the ease of implementation, particularly for high dimensional problems. A numerical example of a eight dimensional Skorohod problem is presented to illustrate the applicability of the approach. The paper [12] based on this chapter (joint work with A. Budhiraja and S. Rubenthaler) is currently under revision for *Mathematics of Operations Research*.

Chapter 2

Preliminaries

The theory of large deviations is concerned with the study of asymptotic exponential decay rate of "rare events". Consider for example, a Poisson process on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with constant rate γ. By law of large numbers, we know that, for each $t > 0$, $\frac{N(nt)}{n} \to \gamma t$, a.s., as $n \to \infty$. This in particular implies that $\mathbb{P}(|\frac{N(nt)}{n} - \gamma t| \geq \epsilon) \to 0$, for $\epsilon > 0$. Theory of large deviations gives us detailed information about the rate of this convergence. In particular, it can be shown that

$$\mathbb{P}(|\frac{N(nt)}{n} - \gamma t| \geq \epsilon) \sim e^{-nI\{[0,\gamma t-\epsilon]\cup[\gamma t+\epsilon,\infty)\}},$$

where $I : \mathbb{R}_+ \to [0, \infty]$ is the, so called, rate function given as (see [33])

$$I(x) = \inf_{g:\int_0^t \gamma g(s)ds=x} \left\{ \int_0^t (g(s)\log g(s) - g(s) + 1)ds \right\},$$

and $I(F) = \inf_{x\in F} I(x)$, for a Borel set $F \subset \mathbb{R}_+$. The infimum is taken over measurable maps $g : [0, t] \to \mathbb{R}_+$. Such an estimate is obtained by establishing the so called "Large Deviation Principle" (LDP) for the collection of random variables $\{\frac{N(nt)}{n}\}_{n\in\mathbb{N}}$, a precise definition of which is given in Section 2.1. A LDP in fact yields similar estimates for a more general family of sets than unions of intervals. One can similarly establish a LDP for the family of $D([0, T] : \mathbb{R}_+)$ valued random variables $\{\bar{N}^n\}_{n\in\mathbb{N}} \equiv \{\frac{N(nt)}{n} : t \in [0, T]\}_{n\in\mathbb{N}}$ with a corresponding rate function. Here and throughout, notation not introduced in the chapter can be found in the list of notation and symbols on page x. A well known result in the theory of large deviations, known as the *Contraction Principle*, says that, if a family of S valued random variables $\{X^n\}$ satisfies a large deviation principle with rate function I, and f is a continuous map from $S \to S'$ (here S and S' are Polish spaces), then $\{f(X^n)\}$ satisfies a large deviation principle with rate function J, where

$$J(y) = \inf\{I(x) : x \in f^{-1}(y)\}, \quad y \in S'.$$

This result along with the LDP for $\bar{N}^n$ tells us that continuous functionals of the scaled Poisson process $\bar{N}^n$ obey a LDP. Many models in probability and stochastic dynamics are built in terms of functionals of Poisson random measures (PRM) and/or Brownian motions (BM). However, frequently the functionals of interest are not continuous. Furthermore in many situations the functionals may depend on the scaling parameters and hence one needs to handle a sequence of functionals. A promising approach based on certain variational formulas to treat large deviation problems for a family of stochastic dynamical systems driven by PRMs and/or Brownian motions, has been initiated in [15]. These variational formulas described in Section 2.2 will be the starting point of our work.

2.1 Large Deviation Principle and Laplace Principle.

Let $\{X^\epsilon, \epsilon > 0\} \equiv \{X^\epsilon\}$ be a family of random variables defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and taking values in a Polish space (i.e., a complete separable metric space) $\mathcal{E}$. Denote expectation with respect to $\mathbb{P}$ by $\mathbb{E}$. The theory of large deviations is concerned with events A for which probabilities $\mathbb{P}(X^\epsilon \in A)$ converge to zero exponentially fast as $\epsilon \to 0$. The exponential decay rate of such probabilities is typically expressed in terms of a "rate function" I mapping $\mathcal{E}$ into $[0, \infty]$.

Definition 2.1.1 (Rate function)**.** A function $I : \mathcal{E} \to [0, \infty]$ is called a rate function on $\mathcal{E}$, if for each $M < \infty$ the level set $\{x \in \mathcal{E} : I(x) \leq M\}$ is a compact subset of $\mathcal{E}$. For $A \in \mathcal{B}(\mathcal{E})$, we define $I(A) \doteq \inf_{x \in A} I(x)$.

Definition 2.1.2 (Large deviation principle)**.** Let I be a rate function on $\mathcal{E}$. The sequence $\{X^\epsilon\}$ is said to satisfy the large deviation principle on $\mathcal{E}$, as $\epsilon \to 0$, with rate function I if the following two conditions hold:

1. *Large deviation upper bound.* For each closed subset F of $\mathcal{E}$,
$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{P}(X^\epsilon \in F) \leq -I(F).$$
2. *Large deviation lower bound.* For each open subset G of $\mathcal{E}$,
$$\liminf_{\epsilon\to 0} \epsilon \log \mathbb{P}(X^\epsilon \in G) \geq -I(G).$$

If a sequence of random variables satisfies a large deviation principle with some rate function, then the rate function is unique. In many problems one is interested in obtaining exponential estimates on functions which are more general than indicator functions of closed or open sets. This leads to the study of the Laplace principle.

Definition 2.1.3 (Laplace principle)**.** Let I be a rate function on $\mathcal{E}$. The sequence $\{X^\epsilon\}$ is said to satisfy the Laplace principle upper bound (respectively lower bound) on $\mathcal{E}$, as $\epsilon \to 0$, with rate function I if for all $h \in C_b(\mathcal{E})$,
$$\limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left\{\exp\left[-\frac{1}{\epsilon}h(X^\epsilon)\right]\right\} \leq -\inf_{x\in\mathcal{E}}\{h(x)+I(x)\},$$
and, respectively,
$$\liminf_{\epsilon\to 0} \epsilon \log \mathbb{E}\left\{\exp\left[-\frac{1}{\epsilon}h(X^\epsilon)\right]\right\} \geq -\inf_{x\in\mathcal{E}}\{h(x)+I(x)\}.$$
Laplace principle is said to hold for $\{X^\epsilon\}$ with rate function I if both the Laplace upper and lower bounds are satisfied for every bounded continuous function h.

One of the main results in the theory of large deviations is the equivalence between the Laplace principle and the large deviation principle. For a proof we refer the reader to Section 1.2 of [33].

Theorem 2.1.1. *The family $\{X^\epsilon\}$ satisfies the Laplace principle upper (respectively lower) bound with a rate function I on $\mathcal{E}$ if and only if $\{X^\epsilon\}$ satisfies the large deviation upper (respectively lower) bound for all closed sets (respectively open sets) with the rate function I.*

2.2 Infinite Dimensional Brownian Motions and Poisson Random Measures.

In this work we are primarily interested in large deviation behavior of stochastic dynamical systems driven by a Brownian motion (possibly infinite dimensional) and/or a Poisson random measure (PRM). Key ingredients in our proofs will be certain variational representations for nonlinear functionals of infinite dimensional Brownian motions and general PRM's [14, 15, 18]. Using such variational representations, general large deviation principles have been developed in [15, 18], which will be the starting point of our study. In this section, we summarize the variational representations and large deviation results from [15, 18].

2.2.1 Infinite Dimensional Brownian Motions and a Variational Representation.

Let $\{\beta_i\}_{i=1}^\infty$ be an infinite sequence of independent, standard, one dimensional, $\{\mathscr{F}_t\}$-Brownian motions given on the filtered probability space $(\Omega, \mathscr{F}, \mathbb{P}, \{\mathscr{F}_t\})$. We denote the product space of countably infinite copies of the real line by $\mathbb{R}^\infty$. Endowed with the topology of coordinate-wise convergence, $\mathbb{R}^\infty$ is a Polish space. We will consider the Brownian motions on a fixed finite time interval $[0, T]$. Then $\beta = \{\beta_i\}_{i=1}^\infty$ can be regarded as a random variable with values in the Polish space $\mathcal{C}([0, T] : \mathbb{R}^\infty)$ and it represents the simplest model of an infinite dimensional Brownian motion.

We call a function $f : [0, T] \times \Omega \to \mathbb{R}$ elementary if there exist $a, b \in [0, T], a \leq b$, and a bounded $\{\mathscr{F}_a\}$-measurable random variable X such that $f(s, \omega) = X(\omega)1_{(a,b]}(s)$. A finite sum of elementary functions is referred to as a simple function. We denote by $\bar{\mathscr{S}}$ the class of all simple functions. The predictable σ -field $\mathscr{P}$ on $[0, T] \times \Omega$ is the σ -field generated by $\bar{\mathscr{S}}$. For a Hilbert space $(H, \langle \cdot, \cdot \rangle)$, a function $f : [0, T] \times \Omega \to H$ is called an H valued predictable process if it is $\mathscr{P}/\mathcal{B}(H)$ -measurable. Let $\mathscr{P}_2(H)$ be the family of all H valued predictable processes ϕ such that $\int_0^T ||\phi(s)||^2 ds < \infty$ a.s., where $|| \cdot ||$ is the norm in the Hilbert space H. We denote by l_2 the Hilbert space of real sequences $a = (a_i)$ satisfying $||a||^2 = \sum_{i=1}^{\infty} a_i^2 < \infty$, with the usual inner product. Note that in the case $H = l_2$, $u \in \mathscr{P}_2(H) = \mathscr{P}_2(l_2)$ can be written as $u = \{u_i\}_{i=1}^{\infty}$, where $u_i \in \mathscr{P}_2(\mathbb{R})$ and $\sum_{i=1}^{\infty} \int_0^T |u_i(s)|^2 ds < \infty$ a.s.

The following variational representation for functionals of an infinite dimensional Brownian motion is taken from [15] (see also [14]).

Theorem 2.2.1. *Let $f \in M_b(\mathcal{C}([0, T] : \mathbb{R}^{\infty}))$. Then,*

$$-\log \mathbb{E}(\exp\{-f(\beta)\}) = \inf_{u \in \mathscr{P}_2(l_2)} \mathbb{E}\left(\frac{1}{2}\int_0^T ||u(s)||^2 ds + f\left(\beta + \int_0^{\cdot} u(s) ds\right)\right).$$

This result and its variants have been used in several studies of large deviations for small noise stochastic dynamical systems [5, 15, 17, 23, 30, 31, 61, 63, 69, 71, 74, 81, 86, 89, 90].

2.2.2 Poisson Random Measure and a Variational Representation.

Let $\mathbb{X}$ be a locally compact Polish space. Let $\mathcal{M}_{FC}(\mathbb{X})$ be the space of all measures ν on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ such that $\nu(K) < \infty$ for every compact K in $\mathbb{X}$. Endow $\mathcal{M}_{FC}(\mathbb{X})$ with the weakest topology such that for every $f \in C_c(\mathbb{X})$ (the space of continuous functions with compact support), the function $\nu \mapsto \langle f, \nu \rangle = \int_{\mathbb{X}} f(u) d\nu(u)$, $\nu \in \mathcal{M}_{FC}(\mathbb{X})$ is continuous. This topology can be metrized such that $\mathcal{M}_{FC}(\mathbb{X})$ is a Polish space (see e.g. [18]). Fix $T \in (0, \infty)$ and let $\mathbb{X}_T = [0, T] \times \mathbb{X}$. Fix a measure $\nu \in \mathcal{M}_{FC}(\mathbb{X})$, and let $\nu_T = \lambda_T \otimes \nu$, where λ_T is Lebesgue measure on $[0, T]$.

We recall that a Poisson random measure $\mathbf{n}$ on $\mathbb{X}_T$ with mean measure (or intensity measure) ν_T is a $\mathcal{M}_{FC}(\mathbb{X}_T)$ valued random variable such that for each $B \in \mathcal{B}(\mathbb{X}_T)$ with $\nu_T(B) < \infty$, $\mathbf{n}(B)$ is Poisson distributed with mean $\nu_T(B)$ and for disjoint $B_1, ..., B_k \in \mathcal{B}(\mathbb{X}_T)$, $\mathbf{n}(B_1), ..., \mathbf{n}(B_k)$ are mutually independent random variables (cf. [49]). Denote by $\mathbb{P}$ the measure induced by $\mathbf{n}$ on $(\mathcal{M}_{FC}(\mathbb{X}_T), \mathcal{B}(\mathcal{M}_{FC}(\mathbb{X}_T)))$. Then letting $\mathbb{M} = \mathcal{M}_{FC}(\mathbb{X}_T)$, $\mathbb{P}$ is the unique probability measure on $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ under which the canonical map, $N : \mathbb{M} \to \mathbb{M}, N(m) \doteq m$, is a Poisson random measure with intensity measure ν_T. With applications to large deviations in mind, we also consider, for $\theta > 0$, probability measures $\mathbb{P}_\theta$ on $(\mathbb{M}, \mathcal{B}(\mathbb{M}))$ under which N is a Poisson random measure with intensity $\theta \nu_T$. The corresponding expectation operators will be denoted by $\mathbb{E}$ and $\mathbb{E}_\theta$, respectively. We now present a variational representation, obtained in [18], for $-\log \mathbb{E}_\theta(\exp[-F(N)])$, where $F \in M_b(\mathbb{M})$, in terms of a Poisson random measure constructed on a larger space. We begin by describing this construction.

Sometimes, the analysis of large deviation properties for a process is simplified considerably by a convenient control representation for the exponential integrals appearing in the Laplace principle. In contrast with the case of Brownian motion, the formulation of a useful representation is not immediate for Poisson noise. With a Poisson random measure, one needs a control that alters the intensity at time t and for jump type x from that of the underlying PRM to essentially any value in $[0, \infty)$ in a non-anticipating fashion. To accommodate this form of control, we augment the space of jump times and jump types by a variable $r \in [0, \infty)$, and consider in place of the original PRM one whose intensity is a product of ν_T and Lebesgue measure on

r. The desired jump intensities can then be obtained by "thinning" this variable.

More precisely, we let $\mathbb{Y} = \mathbb{X} \times [0, \infty)$ and $\mathbb{Y}_T = [0, T] \times \mathbb{Y}$. Let $\bar{\mathbb{M}} = \mathcal{M}_{FC}(\mathbb{Y}_T)$ and let $\bar{\mathbb{P}}$ be the unique probability measure on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$ under which the canonical map, $\bar{N} : \bar{\mathbb{M}} \to \bar{\mathbb{M}}, \bar{N}(m) \doteq m$, is a Poisson random measure with intensity measure $\bar{\nu}_T = \lambda_T \otimes \nu \otimes \lambda_\infty$, with λ_∞ Lebesgue measure on $[0, \infty)$. The corresponding expectation operator will be denoted by $\bar{\mathbb{E}}$. Let $\mathcal{F}_t \doteq \sigma\{\bar{N}((0, s] \times A) : 0 \leq s \leq t, A \in \mathcal{B}(\mathbb{Y})\}$, and let $\bar{\mathcal{F}}_t$ denote the completion under $\bar{\mathbb{P}}$. We denote by $\bar{\mathcal{P}}$ the predictable σ-field on $[0, T] \times \bar{\mathbb{M}}$ with the filtration $\{\bar{\mathcal{F}}_t : 0 \leq t \leq T\}$ on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}))$. Let $\bar{\mathcal{A}}$ be the class of all $(\bar{\mathcal{P}} \otimes \mathcal{B}(\mathbb{X}))/\mathcal{B}[0, \infty)$-measurable maps $\varphi : \mathbb{X}_T \times \bar{\mathbb{M}} \to [0, \infty)$. For $\varphi \in \bar{\mathcal{A}}$, define a counting process N^φ on $\mathbb{X}_T$ by

$$N^\varphi((0, t] \times U) = \int_{(0,t] \times U} \int_{(0,\infty)} 1_{[0,\varphi(s,x)]}(r) \bar{N}(dsdxdr), \quad t \in [0, T], U \in \mathcal{B}(\mathbb{X}). \tag{2.2.1}$$

N^φ is then the controlled random measure, with φ selecting the intensity for the points at location x and time s, in a possibly random but non-anticipating way. When $\varphi(s, x, \bar{m}) \equiv \theta \in (0, \infty)$, we write $N^\varphi = N^\theta$. Note that N^θ has the same distribution with respect to $\bar{\mathbb{P}}$ as N has with respect to $\mathbb{P}_\theta$. Define $l : [0, \infty) \to [0, \infty)$ by

$$l(r) = r \log r - r + 1, \quad r \in [0, \infty).$$

For any $\varphi \in \bar{\mathcal{A}}$ the quantity

$$L_T(\varphi) = \int_{\mathbb{X}_T} l(\varphi(t, x, \omega)) \nu_T(dtdx) \tag{2.2.2}$$

is well defined as a $[0, \infty]$-valued random variable. The following is a representation formula proved in [18].

Theorem 2.2.2. *Let $F \in M_b(\mathbb{M})$. Then, for $\theta > 0$,*

$$-\log \mathbb{E}_\theta(e^{-F(N)}) = -\log \bar{\mathbb{E}}(e^{-F(N^\theta)}) = \inf_{\varphi \in \bar{\mathcal{A}}} \bar{\mathbb{E}} \left[\theta L_T(\varphi) + F(N^{\theta\varphi})\right].$$

2.2.3 Representations for Functionals of both a PRM and an Infinite dimensional BM.

One can combine Theorems 2.2.1 and 2.2.2 to obtain variational representations for a joint functional of a BM and a PRM. We begin by constructing a suitable probability space. As before, denote the product space of countable infinite copies of the real line by $\mathbb{R}^\infty$, and recall that with the topology of coordinate-wise convergence $\mathbb{R}^\infty$ is a Polish space. We denote the Polish space $\mathcal{C}([0,T]:\mathbb{R}^\infty)$ by $\mathbb{W}$ and denote by $\mathbb{V}$ the product space $\mathbb{W}\times\mathbb{M}$. Let $\bar{\mathbb{V}}=\mathbb{W}\times\bar{\mathbb{M}}$. Abusing notation from Section 2.2.2, let $N:\mathbb{V}\to\mathbb{M}$ be defined by $N(w,m)=m$, for $(w,m)\in\mathbb{V}$. The map $\bar{N}:\bar{\mathbb{V}}\to\bar{\mathbb{M}}$ is defined analogously. Let $\beta=\{\beta_i\}_{i=1}^\infty$ be coordinate maps on $\mathbb{V}$ defined as $\beta_i(w,m)=w_i$. Analogous maps on $\bar{\mathbb{V}}$ are denoted again as $\beta=\{\beta_i\}_{i=1}^\infty$. Define $\mathcal{G}_t\doteq\sigma\{N((0,s]\times A),\beta_i(s):0\le s\le t, A\in\mathcal{B}(\mathbb{X}), i\ge 1\}$. For $\theta>0$, denote by $\mathbb{P}_\theta$ the unique probability measure on $(\mathbb{V},\mathcal{B}(\mathbb{V}))$ such that under $\mathbb{P}_\theta$:

1. $\{\beta_i\}_{i=1}^\infty$ is an i.i.d. family of standard Brownian motions.
2. N is a PRM with intensity measure $\theta\nu_T$.
3. $\{\beta_i(t),t\in[0,T]\}$, $\{N((0,t]\times A),t\in[0,T]\}$ are $\mathcal{G}_t$-martingales for every $i\ge 1$, and $A\in\mathcal{B}(\mathbb{X})$ with $\nu(A)<\infty$.

Define $(\bar{\mathbb{P}},\{\bar{\mathcal{G}}_t\})$ on $(\bar{\mathbb{V}},\mathcal{B}(\bar{\mathbb{V}}))$ analogous to $(\mathbb{P}_\theta,\{\mathcal{G}_t\})$ by replacing $(N,\theta\nu_T)$ with $(\bar{N},\bar{\nu}_T)$ in the above. Throughout we will consider the $\bar{\mathbb{P}}$-completion of the filtration $\{\bar{\mathcal{G}}_t\}$ and denote it by $\{\bar{\mathscr{F}}_t\}$. We denote by $\bar{\mathcal{P}}$ the predictable σ-field on $[0,T]\times\bar{\mathbb{V}}$ with the filtration $\{\bar{\mathscr{F}}_t:0\le t\le T\}$ on $(\bar{\mathbb{V}},\mathcal{B}(\bar{\mathbb{V}}))$. Let $\bar{\mathcal{A}}$ be the class of all $(\bar{\mathcal{P}}\otimes\mathcal{B}(\mathbb{X}))/\mathcal{B}[0,\infty)$ measurable maps $\varphi:\mathbb{X}_T\times\bar{\mathbb{V}}\to[0,\infty)$. For $\varphi\in\bar{\mathcal{A}}$, define $L_T(\varphi)$ and the counting process N^φ on $\mathbb{X}_T$ as in Section 2.2.2.

As in Section 2.2.1, we define $\mathscr{P}_2(l_2) \equiv \mathscr{P}_2$ as

$$\mathscr{P}_2 = \left\{ \psi = \{\psi_i\}_{i=1}^{\infty} : \psi_i \text{ is } \bar{\mathcal{P}}/\mathcal{B}(\mathbb{R}) \text{ measurable and } \int_0^T ||\psi(s)||^2 ds < \infty, \text{a.s.}\bar{\mathbb{P}} \right\},$$

and set $\bar{\mathcal{U}} = \mathscr{P}_2 \times \bar{\mathcal{A}}$. For $\psi \in \mathscr{P}_2$ define $\tilde{L}_T(\psi) = \frac{1}{2}\int_0^T ||\psi(s)||^2 ds$ and for $u = (\psi, \varphi) \in \bar{\mathcal{U}}$, set $\bar{L}_T(u) = L_T(\varphi) + \tilde{L}_T(\psi)$. For $\psi \in \mathscr{P}_2$, let $\beta^{\psi} = (\beta_i^{\psi})$ be defined as $\beta_i^{\psi}(t) = \beta_i(t) + \int_0^t \psi_i(s)ds$, $t \in [0,T]$, $i \in \mathbb{N}$. The following result is taken from [18].

Theorem 2.2.3. *Let $F \in M_b(\mathbb{V})$. Then, for $\theta > 0$,*

$$-\log \mathbb{E}_{\theta}(e^{-F(\beta,N)}) = -\log \bar{\mathbb{E}}(e^{-F(\beta,N^{\theta})}) = \inf_{u=(\psi,\varphi)\in\bar{\mathcal{U}}} \bar{\mathbb{E}}\left[\theta \bar{L}_T(u) + F(\beta^{\sqrt{\theta}\psi}, N^{\theta\varphi})\right].$$

2.3 Some General Large Deviation Results.

In this section, we summarize the main large deviation result of [18] that is obtained as a consequence of Theorem 2.2.3.

Let $\{\mathcal{G}^{\epsilon}\}_{\epsilon>0}$ be a family of measurable maps from $\mathbb{V}$ to $\mathbb{U}$, where $\mathbb{U}$ is some Polish space. We will present below a sufficient condition for large deviation principle to hold for the family $Z^{\epsilon} = \mathcal{G}^{\epsilon}\left(\sqrt{\epsilon}\beta, \epsilon N^{\epsilon^{-1}}\right)$, as $\epsilon \to 0$. Define

$$\tilde{S}^N = \left\{ f \in L^2([0,T] : l_2) : \tilde{L}_T(f) \leq N \right\}.$$

Then $\tilde{S}^N$ is a compact subset of $L^2([0,T] : l_2)$ with the weak topology on the Hilbert space. We will throughout consider $\tilde{S}^N$ endowed with this topology. Also, let

$$S^N = \{g : \mathbb{X}_T \to [0,\infty) : L_T(g) \leq N\}. \tag{2.3.1}$$

A function $g \in S^N$ can be identified with a measure $\nu_T^g \in \mathbb{M}$, defined by $\nu_T^g(A) = \int_A g(s,x)\nu_T(dsdx)$, $A \in \mathcal{B}(\mathbb{X}_T)$. This identification induces a topology on S^N, under which S^N is a compact space. The latter is essentially a consequence of the superlinear growth of l and the lower semi-continuity property of relative entropy, as is shown in the following lemma.

Lemma 2.3.1. *For every $N \in \mathbb{N}$, $\{\nu_T^g : g \in S^N\}$ is a compact subset of* $\mathbb{M}$.

Proof. One way to metrize the topology on $\mathbb{M}$, described in Section 2.2.2 (making $\mathbb{M}$ a Polish space), is the following. Consider a sequence of open sets $\{O_j, j \in \mathbb{N}\}$ such that $\bar{O}_j \subset O_{j+1}$, each $\bar{O}_j$ is compact, and $\cup_{j=1}^{\infty} O_j = \mathbb{X}_T$ (cf. Theorem 9.5.21 of [72]). Let $\phi_j(x) = [1 - d(x, O_j)] \vee 0$, where d denotes the metric on $\mathbb{X}_T$. Given any $\mu \in \mathbb{M}$, let $\mu^{(j)} \in \mathbb{M}$ be defined by $\left[d\mu^{(j)}/d\mu\right](x) = \phi_j(x)$. Given $\mu, \nu \in \mathbb{M}$, let

$$\bar{d}(\mu, \nu) = \sum_{j=1}^{\infty} 2^{-j} \left\| \mu^{(j)} - \nu^{(j)} \right\|_{BL},$$

where $\|\cdot\|_{BL}$ denotes the bounded, Lipschitz norm on $\mathcal{M}_F(\mathbb{X}_T)$:

$$\begin{aligned} &\| \ \mu^{(j)} - \nu^{(j)} \Big\|_{BL} \\ &\quad = \sup \left\{ \int_{\mathbb{X}_T} f d\mu^{(j)} - \int_{\mathbb{X}_T} f d\nu^{(j)} : |f|_{\infty} \leq 1, |f(x) - f(y)| \leq d(x, y) \text{ for all } x, y \in \mathbb{X}_T \right\}. \end{aligned}$$

It is straightforward to check that $\bar{d}(\mu, \nu)$ defines a metric under which $\mathbb{M}$ is a Polish space, and that convergence in this metric is essentially equivalent to weak convergence on each compact subset of $\mathbb{X}_T$. Specifically, $\bar{d}(\mu_n, \mu) \to 0$ if and only if for each $j \in \mathbb{N}$, $\mu_n^{(j)} \to \mu^{(j)}$ in the weak topology as finite nonnegative measures, i.e., for all $f \in C_b(\mathbb{X}_T)$

$$\int_{\mathbb{X}_T} f d\mu_n^{(j)} \to \int_{\mathbb{X}_T} f d\mu^{(j)}.$$

Let $\mu_n = \nu_T^{g_n}$. We first show that $\{\mu_n\} \subset \mathbb{M}$ is relatively compact for any sequence $\{g_n\} \subset S^N$. For this, by using a diagonalization method, it suffices to show that $\{\mu_n^{(j)}\} \subset \mathbb{M}$ is relative compact for every j. Next, since $\mu_n^{(j)}$ are supported on a compact subset of $\mathbb{X}_T$ given as $K^j = \overline{\{x | \phi_j(x) \neq 0\}}$, to show $\{\mu_n^{(j)}\} \subset \mathbb{M}$ is relative compact it suffices to show $\sup_n \mu_n^{(j)}(\mathbb{X}_T) < \infty$. The last property will follow from $L_T(g_n) \leq N$, for all n, and the super-linear growth of l. Specifically, let $c \in (0, \infty)$

be such that $z \le c(l(z)+1)$ for all $z \in [0,\infty)$. Then

$$\begin{aligned}\sup_n \mu_n^{(j)}(\mathbb{X}_T) &= \sup_n \int_{\mathbb{X}_T} \phi_j(v) g_n(v) \nu_T(dv) \\ &\le \sup_n \int_{K^j} g_n(v)\nu_T(dv) \le c(N + \nu_T(K^j)) < \infty.\end{aligned}$$

Next, suppose that along a subsequence (without loss of generality, also denoted by $\{\mu_n\}$), $\mu_n \to \mu$. We would like to show that μ is of the form ν_T^g, where $g \in S^N$. For this we will use the lower semi-continuity property of relative entropy. The result holds trivially if $\mu = 0$. Suppose now $\mu \neq 0$. Then there exists $j_0 \in \mathbb{N}$ such that for all $j \ge j_0$, $\inf_{n\in\mathbb{N}} \nu_T^{g_n}(\bar{O}_j) > 0$. For $j \ge j_0$, define

$$c^j = \nu_T^{(j)}(\mathbb{X}_T), \quad \bar{\nu}_T^j = \nu_T^{(j)}/c^j;$$

$$c_n^j = \mu_n^{(j)}(\mathbb{X}_T), \quad \bar{\mu}_n^j = \mu_n^{(j)}/c_n^j;$$

$$c_\mu^j = \mu^{(j)}(\mathbb{X}_T), \quad \bar{\mu}^j = \mu^{(j)}/c_\mu^j.$$

Then $\bar{\nu}_T^j$, $\bar{\mu}_n^j$ and $\bar{\mu}^j$ are probability measures, and

$$\begin{aligned}R(\bar{\mu}_n^j || \bar{\nu}_T^j) &= \frac{1}{c_n^j}\int_{\mathbb{X}_T}\left[\log(g_n(v)) + \log(\frac{c^j}{c_n^j})\right] g_n(v)\phi_j(v)\nu_T(dv) \\ &= \frac{1}{c_n^j}\int_{\mathbb{X}_T}\left[l(g_n(v)) + g_n(v) - 1\right]\phi_j(v)\nu_T(dv) + \log(\frac{c^j}{c_n^j}) \\ &\le \frac{1}{c_n^j}N + 1 - \frac{c^j}{c_n^j} + \log(\frac{c^j}{c_n^j}).\end{aligned}$$

Since $\mu_n^{(j)} \to \mu^{(j)}$, we have $c_n^j \to c_\mu^j$. Thus by the lower semi-continuity property of relative entropy,

$$\begin{aligned}R(\bar{\mu}^j || \bar{\nu}_T^j) &\le \liminf_{n\to\infty} R(\bar{\mu}_n^j || \bar{\nu}_T^j) \\ &\le \liminf_{n\to\infty} \frac{1}{c_n^j}N + 1 - \frac{c^j}{c_n^j} + \log(\frac{c^j}{c_n^j}) \\ &\le \frac{1}{c_\mu^j}N + 1 - \frac{c^j}{c_\mu^j} + \log(\frac{c^j}{c_\mu^j}) < \infty \qquad (2.3.2)\end{aligned}$$

Thus $\mu^{(j)}$ is absolutely continuous with respect to $\nu_T^{(j)}$. Define $g^j = d\mu^{(j)}/d\nu_T^{(j)}$, and $g = g^j$ on $\bar{O}_j$. It is easily checked that g is defined consistently, and that $\mu = \nu_T^g$. Also by a direct calculation,

$$R(\bar{\mu}^{(j)}||\bar{\nu}_T^{(j)}) = \frac{1}{c_\mu^j}\int_{\mathbb{X}_T} l(g(v))\phi_j(v)\nu_T(dv) + 1 - \frac{c^j}{c_\mu^j} + \log(\frac{c^j}{c_\mu^j}).$$

Combining the last display with (2.3.2), we have $\int_{\mathbb{X}_T} l(g(v))\phi_j(v)\nu_T(dv) \le N$, for all j. Sending $j \to \infty$, we see that $g \in S^N$. The result follows. $\square$

Throughout we consider the topology on S^N obtained through the above identification making it a compact space. Let $\bar{S}^N = \tilde{S}^N \times S^N$ with the usual product topology. Let

$$\bar{\mathcal{U}}^N = \{u = (\psi, \varphi) \in \bar{\mathcal{U}} : u(w) \in \bar{S}^N, \bar{\mathbb{P}}\ a.e.w\}.$$

Finally, let $\bar{\mathbb{S}} = \bigcup_{N\ge 1} \bar{S}^N$ and $\mathbb{S} = \bigcup_{N\ge 1} S^N$. The following is the main condition for a large deviation property to hold for $Z^\epsilon = \mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta, \epsilon N^{\epsilon^{-1}}\right)$.

Condition 2.3.1. *There exists a measurable map $\mathcal{G}^0 : \mathbb{V} \to \mathbb{U}$ such that the following hold.*

1. *For $N \in \mathbb{N}$, let $(f_n, g_n), (f, g) \in \bar{S}^N$ be such that $(f_n, g_n) \to (f, g)$ as $n \to \infty$. Then*

$$\mathcal{G}^0\left(\int_0^\cdot f_n(s)ds, \nu_T^{g_n}\right) \to \mathcal{G}^0\left(\int_0^\cdot f(s)ds, \nu_T^g\right).$$

2. *For $N \in \mathbb{N}$, let $u_\epsilon = (\psi_\epsilon, \varphi_\epsilon)$, $u = (\psi, \varphi) \in \bar{\mathcal{U}}^N$ be such that, as $\epsilon \to 0$, u_ϵ converges in distribution to u. Then*

$$\mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta + \int_0^\cdot \psi_\epsilon(s)ds, \epsilon N^{\epsilon^{-1}\varphi_\epsilon}\right) \Rightarrow \mathcal{G}^0\left(\int_0^\cdot \psi(s)ds, \nu_T^\varphi\right).$$

For $\phi \in \mathbb{U}$, let $\bar{\mathbb{S}}_\phi = \left\{(f, g) \in \bar{\mathbb{S}} : \phi = \mathcal{G}^0(\int_0^\cdot f(s)ds, \nu_T^g)\right\}$. Let $\bar{I} : \mathbb{U} \to [0, \infty]$ be defined by

$$\bar{I}(\phi) = \inf_{q=(f,g)\in\bar{\mathbb{S}}_\phi} \left\{\bar{L}_T(q)\right\}, \quad \phi \in \mathbb{U}. \tag{2.3.3}$$

The following theorem from [18] gives the large deviation principle for $\{Z^\epsilon\}_{\epsilon>0}$ under Condition 2.3.1.

Theorem 2.3.1. *For $\epsilon > 0$, let Z^ϵ be defined as $Z^\epsilon = \mathcal{G}^\epsilon\left(\sqrt{\epsilon}\beta, \epsilon N^{\epsilon^{-1}}\right)$ and suppose that Condition 2.3.1 holds. Then $\bar{I}$ given in (2.3.3) is a rate function on $\mathbb{U}$ and the family $\{Z^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle, as $\epsilon \to 0$, with rate function $\bar{I}$.*

Large deviation results of functionals that depend only on a PRM follow as a special case of Theorem 2.3.1. For future reference we record this special case below. Let

$$\mathcal{U}^N = \{\varphi \in \bar{\mathcal{A}} : \varphi(w) \in S^N, \mathbb{P}\ a.e.w\}.$$

Let $\{\mathcal{G}^\epsilon\}_{\epsilon>0}$ be a family of measurable maps from $\mathbb{M}$ to $\mathbb{U}$. The following is the analogue of Condition 2.3.1.

Condition 2.3.2. *There exists a measurable map $\mathcal{G}^0 : \mathbb{M} \to \mathbb{U}$ such that the following hold.*

1. *For $N \in \mathbb{N}$, let $g_n, g \in S^N$ be such that $g_n \to g$ as $n \to \infty$. Then*

$$\mathcal{G}^0\left(\nu_T^{g_n}\right) \to \mathcal{G}^0\left(\nu_T^{g}\right).$$

2. *For $N \in \mathbb{N}$, let $\varphi_\epsilon, \varphi \in \mathcal{U}^N$ be such that, as $\epsilon \to 0$, φ_ϵ converges in distribution to φ. Then*

$$\mathcal{G}^\epsilon\left(\epsilon N^{\epsilon^{-1}\varphi_\epsilon}\right) \Rightarrow \mathcal{G}^0\left(\nu_T^{\varphi}\right).$$

For $\phi \in \mathbb{U}$, define $\mathbb{S}_\phi = \{g \in \mathbb{S} : \phi = \mathcal{G}^0(\nu_T^g)\}$. Let $I : \mathbb{U} \to [0, \infty]$ be defined by

$$I(\phi) = \inf_{g \in \mathbb{S}_\phi} \{L_T(g)\}, \quad \phi \in \mathbb{U}. \tag{2.3.4}$$

The following theorem is an immediate consequence of Theorem 2.3.1.

Theorem 2.3.2. *For $\epsilon > 0$, let Z^ϵ be defined as $Z^\epsilon = \mathcal{G}^\epsilon\left(\epsilon N^{\epsilon^{-1}}\right)$, and suppose that Condition 2.3.2 holds. Then I defined as in (2.3.4) is a rate function on $\mathbb{U}$ and the family $\{Z^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle, as $\epsilon \to 0$, with rate function I.*

Chapter 3

A Stochastic Partial Differential Equation Model for Spread of a Pollutant

3.1 Introduction.

In hydrology literature (see [43] for example), ordinary differential equations (ODEs) of the following type are often used to model the spread of a pollutant in a river or air, or the water quality in a basin or reservoir:

$$D\triangle\phi - V \cdot \nabla\phi - \alpha\phi + Q = 0.$$

Here $\phi(x)$ represents the water quality or chemical concentration at location x; $\triangle$ is the Laplacian operator modeling the diffusion of the chemical; D is the coefficient capturing the strength of the diffusion effect. The term $V \cdot \nabla\phi$ models the convection term, here ∇ is the gradient operator and V is the velocity vector. The scalar $\alpha \geq 0$ can be interpreted as the rate of dissipation of the chemical and $Q \geq 0$ is the "load" or pollutant issued from outside.

The above deterministic equation models the steady state density profile of the pollutant and does not take into account any temporal or stochastic variability. A dynamic stochastic model for pollutant spread described through a stochastic partial differential equation (SPDE) driven by a PRM was studied in [52]. We begin by describing this model in a one dimensional setting, where it describes the evolution of a pollutant deposited at different sites along a river. We will then present some time asymptotic results that describe the long term profile of the pollutant through certain law of large number (LLN) type results. Our goal in this work is to study probabilities of deviations from the nominal behavior, described through the LLN, by establishing a suitable large deviation principle. Theorem 2.3.2 will be a crucial ingredient in our proofs.

3.2 Dynamic SPDE Model.

The model considered here describes spread of concentration of undesired chemicals which are released by several different sources along a river. Suppose that there are r such sources located at different sites κ_1, ...,$\kappa_r \in [0,l]$, where the interval $[0,l]$ represents the river. These sources release chemicals according to independent Poisson streams $N_i(t)$, with rate f_i, $i = 1, ..., r$, and with random magnitudes $A_i^j(\omega)$, $j \in \mathbb{N}$, $i = 1, ...r$, which are mutually independent with magnitudes in the i^{th} stream having common distribution $F_i(da)$. The chemicals released by the i^{th} stream are deposited uniformly over $(\kappa_i - \epsilon_i, \kappa_i + \epsilon_i), i = 1, ..., r$. We assume without loss of generality that $\cup_{i=1}^r (\kappa_i - \epsilon_i, \kappa_i + \epsilon_i) \subseteq [0,l]$.

Formally, the model describing the evolution of concentration is written as follows:

$$\begin{aligned} \frac{\partial}{\partial t} u(t,x) =& D\frac{\partial^2}{\partial x^2} u(t,x) - V\frac{\partial}{\partial x} u(t,x) - \alpha u(t,x) \\ &+ \sum_{i=1}^{r} \sum_{j} \frac{A_i^j(\omega)}{2\epsilon_i} 1_{(\kappa_i-\epsilon_i,\kappa_i+\epsilon_i)}(x) 1_{t=\tau_i^j(\omega)}, \end{aligned} \tag{3.2.1}$$

where $\tau_i^j(\omega)$, $j \in \mathbb{N}$ are the jump times of N_i. The equation is considered with a Neumann boundary condition on $[0,l]$.

The above equation can be regarded as a stochastic partial differential equation driven by a Poisson random measure with solutions in the dual of a suitable nuclear space. The Poisson random measure N driving the equation is a random measure on the space $\mathbb{R}_+ \times [0,l] \times \mathbb{R}_+$ given as

$$N([0,t] \times A \times B) = \sum_{i=1}^{r} 1_A(\kappa_i) \sum_{j=1}^{N_i(t)} 1_B(A_i^j(\omega)), \quad t \geq 0,\ A \in \mathcal{B}([0,l]),\ B \in \mathcal{B}(\mathbb{R}_+),$$

and the intensity measure of N is given as

$$\nu_0([0,t] \times A \times B) = t \sum_{i=1}^{r} 1_A(\kappa_i) f_i F_i(B), \qquad A \in \mathcal{B}([0,l]),\ B \in \mathcal{B}(\mathbb{R}_+),\ t \geq 0. \quad (3.2.2)$$

We now describe the solution space for the equation in (3.2.1). We begin with some basic definitions (see [52]).

Definition 3.2.1. Let $\mathcal{E}$ be a vector space. A family of norms $\{||\cdot||_p : p \in \mathbb{N}_0\}$ on $\mathcal{E}$ is called **compatible** if for any $p, q \in \mathbb{N}_0$, whenever $\{x_n\} \subseteq \mathcal{E}$ is a Cauchy sequence with respect to both $||\cdot||_p$ and $||\cdot||_q$, and converges to 0 with respect to one norm, then it also converges to 0 with respect to the other norm. The family is said to be **increasing** if for all $x \in \mathcal{E}$, $||x||_p \leq ||x||_q$ whenever $p \leq q$.

Definition 3.2.2. A separable Frèchet space Φ is called a **countable Hilbertian space** if its topology is given by an increasing sequence $||\cdot||_n$, $n \in \mathbb{N}_0$, of compatible Hilbertian norms. A countable Hilbertian space Φ is called **nuclear** if for each $n \in \mathbb{N}_0$ there exists $m > n$ such that the canonical injection from Φ_m into Φ_n is Hilbert-Schmidt, where Φ_k, for each $k \in \mathbb{N}_0$, is the completion of Φ with respect to $||\cdot||_k$.

If Φ, $\{\Phi_n\}_{n\in\mathbb{N}_0}$ are as above, then $\{\Phi_n\}_{n\in\mathbb{N}_0}$ is a sequence of decreasing Hilbert spaces and $\Phi = \cap_{n=0}^{\infty}\Phi_n$. Identify Φ_0' with Φ_0 using Riesz's representation theorem, and denote Φ_n' (the dual of Φ_n) by Φ_{-n} and corresponding norm by $||\cdot||_{-n}$, $n \in \mathbb{N}_0$. Then $\{\Phi_{-n}\}_{n\in\mathbb{N}_0}$ is a sequence of increasing Hilbert spaces, Φ' (the dual of Φ) is sequentially complete and $\Phi' = \cup_{n=0}^{\infty}\Phi_{-n}$ (see Theorem 1.3.1 of [52]).

A natural Countable Hilbertian Nuclear Space (CHNS) associated with equation (3.2.1) is described as follows (see [52]).

Let $\rho \in \mathcal{M}_F[0,l]$ be defined as

$$\rho(A) = \int_A e^{-2cx} dx; \quad A \in \mathcal{B}[0,l],$$

where $c = \frac{V}{2D}$. Let $H = L^2([0,l],\rho)$. Then $\{\phi_j\}_{j\in\mathbb{N}_0}$ defined below defines a complete orthonormal system on H of eigen-functions of L which is defined as

$$L = D\frac{\partial^2}{\partial x^2} - V\frac{\partial}{\partial x}, \tag{3.2.3}$$

with Neumann boundary condition with corresponding eigenvalues denoted by $\{-\lambda_j\}_{j\in\mathbb{N}_0}$.

$$\phi_0(x) = \sqrt{\frac{2c}{1-e^{-2cl}}}, \quad \phi_j(x) = \sqrt{\frac{2}{l}} e^{cx} \sin\left(\frac{j\pi}{l}x + \alpha_j\right);$$

$$\alpha_j = \tan^{-1}(-\frac{j\pi}{lc}), \quad j = 1,2,...;$$

$$\lambda_0 = 0, \quad \lambda_j = D\left(c^2 + \left(\frac{j\pi}{l}\right)^2\right).$$

For $\phi \in H$ and $n \in \mathbb{Z}$, let

$$||\phi||_n^2 = \sum_{j=0}^{\infty} \langle \phi, \phi_j \rangle^2 (1+\lambda_j)^{2n},$$

where $\langle \phi, \psi \rangle$ is the inner product on H. Define

$$\Phi = \{\phi \in H : ||\phi||_n < \infty, \forall n \in \mathbb{Z}\},$$

and let Φ_n be the completion of Φ with respect to the norm $||\cdot||_n$. Note that $\Phi_0 = H$, and it can be checked that Φ is a CHNS.

We consider SDEs driven by PRMs with solutions in duals of nuclear spaces. Namely, we consider a SDE of the form

$$X_t = X_0 + \int_0^t A(s, X_s) ds + \int_0^t \int_{\mathbb{X}} G(s, X_{s-}, v) \tilde{N}(dsdv) \tag{3.2.4}$$

on the dual Φ' of a CHNS Φ, where $A : \mathbb{R}_+ \times \Phi' \to \Phi'$, $G : \mathbb{R}_+ \times \Phi' \times \mathbb{X} \to \Phi'$ are measurable maps, $\mathbb{X}$ is a locally compact space, $N(dsdv)$ is a PRM on $\mathbb{R}_+ \times \mathbb{X}$ with intensity measure $\lambda \otimes \nu$, λ is the Lebesgue measure and ν is a σ-finite measure on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$, and $\tilde{N}(dsdv)$ is the compensated random measure of N, i.e.

$$\tilde{N}([0,t] \times B) = N([0,t] \times B) - t\nu(B),$$

$\forall B \in \mathcal{B}(\mathbb{X})$ with $\nu(B) < \infty$. Theorem 6.3.1 of [52] gives sufficient conditions on A and G under which (3.2.4) has a unique solution.

For the equation in (3.2.1), $\mathbb{X} = [0,l] \times \mathbb{R}_+$ and the measurable maps A and G do not depend on the first variable and first two variables, respectively. More precisely, $A : \Phi' \to \Phi'$ and $G : [0,l] \times \mathbb{R}_+ \to \Phi'$ are given as

$$A(u)[\phi] = u[L\phi] - \alpha u[\phi] + \sum_{i=1}^{r} \frac{a_i f_i}{2\epsilon_i} \int_{\kappa_i - \epsilon_i}^{\kappa_i + \epsilon_i} \phi(y)\rho(y)dy, \quad \phi \in \Phi, u \in \Phi',$$

$$G(x,a)[\phi] = \begin{cases} \frac{a}{2\epsilon_i} \int_{\kappa_i - \epsilon_i}^{\kappa_i + \epsilon_i} \phi(y)\rho(y)dy & x = \kappa_i, i = 1, \cdots, r \\ 0 & \text{otherwise} \end{cases} ; \; (x,a) \in [0,l] \times \mathbb{R}_+, \; \phi \in \Phi,$$

where $a_i = \int_{\mathbb{R}_+} aF_i(da)$, L is defined as in (3.2.3). Note that the operator $-L$ on H is positive definite and self-adjoint. With this notation, equation (3.2.1) can be written in the form of (3.2.4) as

$$u_t = u_0 + \int_0^t A(u_s)ds + \int_0^t \int_{[0,l] \times \mathbb{R}_+} G(x,a)\tilde{N}(dsdxda). \tag{3.2.5}$$

Using Theorem 6.3.1 of [52], (3.2.5) has a unique Φ'-valued solution. More precisely, let $(\Omega, \mathscr{F}, \mathbb{P}, \{\mathscr{F}_t\})$ be a filtered probability space on which is given a Poisson random measure N with intensity measure ν_0 as in (3.2.2), such that $N([0,t] \times A \times B) - \nu_0([0,t] \times A \times B)$ is a $\{\mathscr{F}_t\}$ martingale for all $A \in \mathcal{B}([0,l])$, $B \in \mathcal{B}(\mathbb{R}_+)$ satisfying $\nu_0([0,t] \times A \times B) < \infty$, $t \geq 0$; and let u_0 be a $\mathscr{F}_0$ measurable Φ' valued random variable. Then there exists a unique $\{\mathscr{F}_t\}$ adapted Φ' valued process $\{u_t\}_{t \geq 0}$ such that $\forall \phi \in \Phi$, the following integral equation is satisfied,

$$u_t[\phi] = u_0[\phi] + \int_0^t A(u_s)[\phi]ds + \int_0^t \int_{[0,l] \times [0,\infty)} G(x,a)[\phi]\tilde{N}(dsdxda). \tag{3.2.6}$$

In fact, for the setting considered here, it follows from Theorem 7.3.1 of [52] that if $u_0 \in H$ a.s., then $u_\cdot \in D([0,\infty), H)$ a.s..

We can also use Green's function to represent the solution. The Green's function of the operator L with Neumann boundary conditions is given by

$$p(t,x,y) = \sum_{j=0}^{\infty} e^{-\lambda_j t}\phi_j(x)\phi_j(y), \quad t \geq 0,\ x, y \in [0,l].$$

In terms of p, the solution $u(t,x)$ can be represented as

$$\begin{aligned} u(t,x) = e^{-\alpha t}\int_0^l u_0(y)p(t,x,y)\rho(y)dy \\ + \int_0^t \int_{[0,l]\times[0,\infty)} e^{-\alpha(t-s)}\frac{a}{2\epsilon_i}1_{y=\kappa_i}\left(\int_{\kappa_i-\epsilon_i}^{\kappa_i+\epsilon_i} p(t-s,x,u)\rho(u)du\right)N(dsdyda). \end{aligned}$$

We set $b_i = \int_{\mathbb{R}_+} a^2 F_i(da)$ and $\psi_j(\kappa_i, \epsilon_i) = \frac{1}{2\epsilon_i}\int_{\kappa_i-\epsilon_i}^{\kappa_i+\epsilon_i} \phi_j(y)\rho(y)dy$ for future use.

We are interested in the time asymptotic behavior of u_t. This behavior is quite different for the cases $\alpha = 0$, and $\alpha > 0$. Thus we consider the two cases separately in the following two sections.

3.3 The case $\alpha = 0$.

Throughout this section we will assume that $\alpha = 0$. In this case, limiting behavior of $\frac{1}{t}u_t$, as $t \to \infty$, is studied in [52]. In particular the following result is obtained.

Theorem 3.3.1. *As t tends to infinity,*

$$\frac{1}{t}u_t \to \sum_{i=1}^{r} f_i a_i \psi_0(\kappa_i, \epsilon_i)\phi_0 \text{ in } H, \text{ almost surely.}$$

Note that ϕ_0 is a constant function, thus the above result says that asymptotically, the rate of growth of the concentration at each site is the same and is given explicitly in terms of the mean deposition magnitudes a_i. This suggests that these mean deposition amounts can be used as a basis for regulating chemical levels. However, looking at these values alone can be misleading, and one would like to know the probability of deviations, particularly large deviations, from the nominal values described through the right side of the above expression. Thus it is of interest to study the large deviation behavior of $\frac{1}{t}u_t$ as t becomes large. More specifically, we will establish a large deviation principle for the sequence $\{\frac{u_{nt}}{n} : n \in \mathbb{N}\}$ of H valued random variables, as $n \to \infty$.

For this study, we will use Theorem 2.3.2. For simplicity, we take the initial profile u_0 to be a fixed (non-random) element of H. In order to apply Theorem 2.3.2, we will take the underlying filtered probability space to be the space $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}), \bar{\mathbb{P}}, \bar{\mathscr{F}}_t)$ introduced in Section 2.2.2 with $\mathbb{X} = [0, l] \times [0, \infty)$, and $\nu(dx \times da) = \nu_0([0, 1] \times dx \times da)$. Let, for $n \in \mathbb{N}$, $\eta_n = \frac{1}{n}N^n$, where N^n is as introduced below (2.2.1). From Theorem 7.3.1 of [52], it follows that for each $n \in \mathbb{N}$, there is a unique $D([0, 1], H)$ valued process $\{\hat{u}^n\}$ solving the following stochastic integral equation

$$\begin{aligned} \hat{u}^n_t =& \frac{1}{n} T_{nt} u_0 + \phi_0 \int_0^t \int_{[0,l]\times[0,\infty)} G(x, a)[\phi_0] \eta_n(ds, dx, da) \\ &+ \sum_{j=1}^{\infty} \phi_j \int_0^t \int_{[0,l]\times[0,\infty)} G(x, a)[\phi_j] e^{-\lambda_j n(t-s)} \eta_n(ds, dx, da), \end{aligned} \tag{3.3.1}$$

where for $u \in \Phi$ and $t > 0$, $T_t u \in H$ is defined as $T_t u = e^{tL} u = \sum_{j=0}^{\infty} e^{-\lambda_j t} u[\phi_j]\phi_j$. From unique solvability of (3.2.5) it follows that $\{\hat{u}^n_t : 0 \le t \le 1\}$ has the same distribution as $\{\frac{1}{n} u_{nt} : 0 \le t \le 1\}$ on $D([0, 1] : H)$. Thus we will instead consider the large deviation behavior of $\{\hat{u}^n : n \in \mathbb{N}\}$ as a sequence of $D([0, 1] : H) \equiv \mathbb{U}$ valued random variables.

We will impose the following exponential integrability condition on the magnitude distribution of the pollutant.

Assumption 3.3.1. *There exists $\delta > 0$ such that*

$$\int_0^\infty e^{\delta a^2} F_i(da) < \infty, \qquad \forall i = 1, ..., r.$$

Due to pathwise unique solvability of (3.3.1), one can represent, for each $n \in \mathbb{N}$, $\hat{u}^n = \mathcal{G}^n(\frac{1}{n}N^n)$ for some measurable map $\mathcal{G}^n : \mathbb{M} \to \mathbb{U}$. Define the map $\mathcal{G}^0 : \mathbb{M} \to \mathbb{U}$ as follows. For $\nu \in \mathbb{M}$, let

$$\mathcal{G}^0(\nu) = \phi_0 \int_0^t \int_{[0,l]\times[0,\infty)} G(x,a)[\phi_0]\nu(ds,dx,da),$$

if $\int_{[0,t]\times[0,l]\times[0,\infty)} |G(x,a)[\phi_0]|\nu(ds,dx,da) < \infty$. If the latter integral is infinite, we set $\mathcal{G}^0(\nu) = 0$ (i.e. the zero trajectory in $D([0,1] : H)$). Define I through (2.3.4).

The following is the main result of this section.

Theorem 3.3.2. *Under Assumption 3.3.1, I is a rate function on $\mathbb{U}$ and the family $\{\hat{u}^n\}_{n\in\mathbb{N}}$ satisfies a large deviation principle, as $n \to \infty$, on $D([0,1] : H)$, with rate function I.*

The key step in the proof of Theorem 3.3.2 is the verification of Condition 2.3.2 (with $T = 1$). We first consider Part 1 of the condition.

Part 1 of Condition 2.3.2:

Let $g_n, g \in S^N$ be such that $g_n \to g$. We will like to show that $\mathcal{G}^0\left(\nu_1^{g_n}\right) \to \mathcal{G}^0\left(\nu_1^g\right)$ in $\mathbb{U}$ (recall the definition of ν_1^g from Section 2.3). Here

$$\begin{aligned}
\mathcal{G}^0\left(\nu_1^{g_n}\right)(t) &= \phi_0 \int_0^t \int_0^l \int_0^\infty G(x,a)[\phi_0] g_n(s,x,a)\nu_T(ds,dx,da) \\
&= \phi_0 \sum_{i=1}^r \int_0^t \int_0^\infty \frac{a}{2\epsilon_i}\left(\int_{\kappa_i-\epsilon_i}^{\kappa_i+\epsilon_i} \phi_0(y)\rho(y)dy\right) f_i g_n(s,\kappa_i,a)F_i(da)ds \\
&= \phi_0 \sum_{i=1}^r \psi_0(\kappa_i,\epsilon_i) f_i \int_0^t \int_0^\infty a g_n(s,\kappa_i,a)F_i(da)ds.
\end{aligned}$$

So it suffices to show that

$$\int_0^t \int_0^\infty a g_n(s,\kappa_i,a) F_i(da) ds \to \int_0^t \int_0^\infty a g(s,\kappa_i,a) F_i(da) ds,$$

uniformly in $t \in [0,1]$ for all i. This is shown in the following lemma.

Lemma 3.3.1. *Under Assumption 3.3.1, we have*

$$\int_0^t \int_0^\infty a g_n(s,\kappa_i,a) F_i(da) ds \to \int_0^t \int_0^\infty a g(s,\kappa_i,a) F_i(da) ds,$$

uniformly in $t \in [0,1]$ for all i, as $n \to \infty$.

Proof. Fix $i \in \{1,...,r\}$. Let $\Psi_L : [0,\infty) \to [0,\infty)$ be a continuous and bounded function such that $\Psi_L(a) \le (L \wedge a)$ and

$$\Psi_L(a) = \begin{cases} a & a \in [0,L]; \\ 0 & a \in [L+1,\infty). \end{cases} \tag{3.3.2}$$

Recalling the topology on S^N, we have, for every $t \in [0,1]$,

$$\int_0^t \int_0^\infty \Psi_L(a) g_n(s,\kappa_i,a) F_i(da) ds \to \int_0^t \int_0^\infty \Psi_L(a) g(s,\kappa_i,a) F_i(da) ds. \tag{3.3.3}$$

In order to show that the convergence in (3.3.3) is uniform in $t \in [0,1]$, we will obtain a suitable equicontinuity estimate. Note that

$$\int_0^\infty \int_0^1 l(g_n(s,\kappa_i,a)) ds F_i(da) \le \frac{N}{f_i}, \quad \forall i = 1,...,r. \ \forall n \in \mathbb{N}. \tag{3.3.4}$$

We will use the following inequality. For $a, b \in (0,\infty)$,

$$ab \le e^{\sigma a} + \frac{1}{\sigma}(b \log b - b + 1) = e^{\sigma a} + \frac{1}{\sigma} l(b), \quad \forall \sigma \in (1,\infty). \tag{3.3.5}$$

Fix $0 \le t_0 < t_1 \le 1$. Then, for any $\sigma \in (1,\infty)$,

$$\begin{aligned} \left| \int_{t_0}^{t_1} \int_0^\infty \Psi_L(a) g_n(s,\kappa_i,a) F_i(da) ds \right| &\le L \left| \int_{t_0}^{t_1} \int_0^\infty g_n(s,\kappa_i,a) F_i(da) ds \right| \\ &\le L \left[(t_1 - t_0) e^{\sigma} + \frac{1}{\sigma f_i} N \right], \end{aligned}$$

where the last inequality follows from using (3.3.4) and applying (3.3.5) with $a = 1$ and $b = g_n$. Equicontinuity of

$$[0,1] \ni t \mapsto \int_0^t \int_0^\infty \Psi_L(a) g_n(s, \kappa_i, a) F_i(da) ds$$

is an immediate consequence of the above estimate. Thus the convergence in (3.3.3) holds uniformly in $t \in [0,1]$.

Next, we show that

$$\sup_{0 \le t \le 1} \sup_n \left| \int_0^t \int_0^\infty (\Psi_L(a) - a) g_n(s, \kappa_i, a) F_i(da) ds \right| \to 0, \quad as \quad L \to \infty. \tag{3.3.6}$$

Fix $\epsilon > 0$. Using (3.3.5) once again, the above quantity is bounded above by

$$\begin{aligned}
&\sup_n \int_0^1 \int_{a \ge L} a g_n(s, \kappa_i, a) F_i(da) ds \\
&\le \int_0^1 \int_{a \ge L} e^{\sigma a} F_i(da) ds + \sup_n \frac{1}{\sigma} \int_0^1 \int_0^\infty l(g_n(s, \kappa_i, a)) F_i(da) ds \\
&\le \int_{a \ge L} e^{\sigma a} F_i(da) + \frac{N}{\sigma f_i}.
\end{aligned}$$

Let σ_0 be large enough, so that $\frac{N}{\sigma f_i} < \frac{\epsilon}{2}$ for all $\sigma \ge \sigma_0$. Fix $\sigma \ge \sigma_0$. From Assumption 3.3.1, $\int e^{\sigma a} F_i(da) < \infty$. Thus choosing L_0 sufficiently large, we have $\forall L \ge L_0$, $\int_{a \ge L} e^{\sigma a} F_i(da) < \frac{\epsilon}{2}$. This proves (3.3.6) and the result follows. $\square$

As an immediate consequence of the lemma and calculations preceding it, we have the following result.

Proposition 3.3.1. *Suppose Assumption 3.3.1 holds. Then for every $N \in \mathbb{N}$, and $g_n, g \in S^N$, $n \in \mathbb{N}$, such that $g_n \to g$, as $n \to \infty$, we have $\mathcal{G}^0(\nu_1^{g_n}) \to \mathcal{G}^0(\nu_1^g)$.*

Next, we proceed to verify Part 2 of Condition 2.3.2.

Part 2 of Condition 2.3.2:

Fix $N \in \mathbb{N}$, and let $\varphi_n, \varphi \in \mathcal{U}^N$ be such that $\varphi_n \Rightarrow \varphi$, as $n \to \infty$. We will like to show that $\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right) \Rightarrow \mathcal{G}^0\left(\nu_1^{\varphi}\right)$.

Note that

$$\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right)(t) = \frac{1}{n}T_{nt}u_0 + \phi_0 \int_0^t \int_{[0,l]\times[0,\infty)} G(x,a)[\phi_0]\frac{1}{n}N^{n\varphi_n}(ds,dx,da) + \sum_{j=1}^{\infty} \phi_j \int_0^t \int_{[0,l]\times[0,\infty)} G(x,a)[\phi_j]e^{-\lambda_j n(t-s)}\frac{1}{n}N^{n\varphi_n}(ds,dx,da),$$

and

$$\mathcal{G}^0\left(\nu_1^{\varphi}\right)(t) = \phi_0 \sum_{i=1}^{r} \psi_0(\kappa_i,\epsilon_i) f_i \int_0^t \int_0^{\infty} a\varphi(s,\kappa_i,a)F_i(da)ds.$$

In order to prove the above weak convergence, the following lemma will be used. For $i = 1, ..., r$, $n \in \mathbb{N}$, let $\varphi_n^i(s,a) = \varphi(s,\kappa_i,a)$; $(s,a) \in [0,1]\times[0,\infty)$.

Lemma 3.3.2. *Fix $i \in \{1,...,r\}$. Suppose that for all $c > 0$, $\int_0^{\infty} e^{ca}F_i(da) < \infty$. Then, for $j \geq 1$,*

$$S^n(j) = \sup_{0\leq t\leq 1}\left(\int_0^t \int_0^{\infty} a e^{-\lambda_j n(t-s)}\varphi_n^i(s,a)F_i(da)ds\right) \to 0, \quad a.s.\ as\ n \to \infty.$$

Also $\sup_{j\in\mathbb{N}_0}\sup_{n\in\mathbb{N}} S^n(j) < \infty$, *a.s.*

Proof. From (3.3.5), for each $c > 1$,

$$\begin{aligned} S^n(j) &\leq \int_0^1 \int_0^{\infty} a\varphi_n^i(s,a)F_i(da)ds \\ &\leq \int_0^{\infty} e^{ca}F_i(da) + \frac{1}{c}\int_0^1\int_0^{\infty} l(\varphi_n^i(s,a))F_i(da)ds \\ &\leq \int_0^{\infty} e^{ca}F_i(da) + \frac{N}{cf_i}, \quad \text{a.s.} \end{aligned}$$

This proves the second statement in the lemma.

Recall the truncation function Ψ_L introduced in (3.3.2). Note that

$$\begin{aligned}S^n(j) \le &\sup_{0\le t\le 1}\left(\int_0^t\int_0^\infty \Psi_L(a)e^{-\lambda_j n(t-s)}\varphi_n^i(s,a)F_i(da)ds\right)\\ &+\sup_{0\le t\le 1}\left(\int_0^t\int_{a\ge L} ae^{-\lambda_j n(t-s)}\varphi_n^i(s,a)F_i(da)ds\right)\\ =&S_1^n(j)+S_2^n(j).\end{aligned}$$

The second term can be bounded as follows.

$$\begin{aligned}S_2^n(j) =& \sup_{0\le t\le 1}\left(\int_0^t\int_{a\ge L} ae^{-\lambda_j n(t-s)}\varphi_n^i(s,a)F_i(da)ds\right)\\ \le&\int_{|a|\ge L} e^{pa}F_i(da)+\frac{1}{p}\int_0^t\int_0^\infty l(\varphi_n^i(s,a))F_i(da)ds\\ =&\theta_1(L,p)+\theta_2(p).\end{aligned}$$

Note that $\theta_2(p)\to 0$ as $p\to\infty$, and for each $p>0$, $\theta_1(L,p)\to 0$ as $L\to\infty$.

For the first term, for $c>0$, let $M(c)=\inf_{x\in[c,\infty)}\frac{l(x)}{x}$. Then $M(c)\to\infty$ as $c\to\infty$. Note that, for $j\ge 1$,

$$\begin{aligned}S_1^n(j) \le& \sup_{0\le t\le 1}\left(\int_{[0,t]\times[0,\infty)} 1_{|\varphi_n^i(s,a)|\le c}\Psi_L(a)e^{-\lambda_j n(t-s)}\varphi_n^i(s,a)F_i(da)ds\right)\\ &+\sup_{0\le t\le 1}\left(\int_{[0,t]\times[0,\infty)} 1_{|\varphi_n^i(s,a)|> c}\Psi_L(a)e^{-\lambda_j n(t-s)}\varphi_n^i(s,a)F_i(da)ds\right)\\ \le& Lc\sup_{0\le t\le 1}\left(\int_0^t e^{-\lambda_j n(t-s)}ds\right)+\frac{L}{M(c)}\int_{[0,1]\times[0,\infty)} 1_{|\varphi_n^i(s,a)|>c}l(\varphi_n^i(s,a))F_i(da)ds\\ \le&\frac{Lc}{n\lambda_j}+\frac{LN}{M(c)f_i}.\end{aligned}$$

Thus we have

$$\begin{aligned}S^n(j) \le& S_1^n(j)+S_2^n(j)\\ \le&\frac{Lc}{n\lambda_j}+\frac{LN}{M(c)f_i}+\theta_1(L,p)+\theta_2(p),\quad \forall L,p,c.\end{aligned}$$

Finally, given $\epsilon>0$, choose p, s.t. $\theta_2(p)<\epsilon/4$; then choose L large enough, s.t. $\theta_1(L,p)<\epsilon/4$; then choose M large enough, s.t. $\frac{LN}{Mf_i}<\epsilon/4$; next let c be large enough so that $M(c)>M$. Finally choose N_0 s.t. if $n\ge N_0$, $\frac{Lc}{n\lambda_j}<\epsilon/4$. Then $S^n(j)\le\epsilon$, for all $n\ge N_0$. The result follows. $\square$

We now prove Part 2 of Condition 2.3.2.

Proposition 3.3.2. *Let $\varphi_n, \varphi \in \mathcal{U}^N$ be such that $\varphi_n \Rightarrow \varphi$, as $n \to \infty$. Then under Assumption 3.3.1, we have*

$$\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right) \Rightarrow \mathcal{G}^0\left(\nu_1^\varphi\right).$$

Proof. Write

$$\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right)(t) = \mathcal{T}_0^n(t) + \mathcal{T}_1^n(t) + \mathcal{T}_2^n(t),$$

where, for $t \in [0,1]$,

$$\begin{aligned}
\mathcal{T}_0^n(t) &= \frac{1}{n}T_{nt}u_0,\\
\mathcal{T}_1^n(t) &= \phi_0 \int_0^t \int_{[0,l]\times[0,\infty)} G(x,a)[\phi_0]\frac{1}{n}N^{n\varphi_n}(ds,dx,da),\\
\mathcal{T}_2^n(t) &= \sum_{j=1}^\infty \mathcal{T}_{2j}^n = \sum_{j=1}^\infty \phi_j \int_0^t \int_{[0,l]\times[0,\infty)} G(x,a)[\phi_j]e^{-\lambda_j n(t-s)}\frac{1}{n}N^{n\varphi_n}(ds,dx,da).
\end{aligned}$$

From the contraction property of $\{T_t\}$,

$$||\mathcal{T}_0^n(t)||_0 = ||\frac{1}{n}T_{nt}u_0||_0 \le \frac{1}{n}||u_0||_0 \to 0,$$

as $n \to \infty$. Thus $\mathcal{T}_0^n \to 0$ in $C([0,1]:H)$.

Next,

$$\begin{aligned}
&\mathcal{T}_{2j}^n(t)\\
&= \phi_j \int_0^t \int_{[0,l]\times[0,\infty)} G(x,a)[\phi_j]e^{-\lambda_j n(t-s)}\frac{1}{n}N^{n\varphi_n}(ds,dx,da)\\
&= \phi_j \sum_{i=1}^r \int_0^t \int_{[0,l]\times[0,\infty)} \frac{a}{2\epsilon_i}\left(\int_{\kappa_i-\epsilon_i}^{\kappa_i+\epsilon_i} \phi_j(y)\rho(y)dy\right) 1_{\{\kappa_i\}}(x)e^{-\lambda_j n(t-s)}\frac{1}{n}N^{n\varphi_n}(ds,dx,da)\\
&= \phi_j \sum_{i=1}^r \int_0^t \int_{[0,l]\times[0,\infty)} a\psi_j(\kappa_i,\epsilon_i)1_{\{\kappa_i\}}(x)e^{-\lambda_j n(t-s)}\frac{1}{n}N^{n\varphi_n}(ds,dx,da)\\
&= \sum_{i=1}^r \left(\overline{\mathcal{T}}_{2j}^{n,i} + \widetilde{\mathcal{T}}_{2j}^{n,i}\right),
\end{aligned}$$

where

$$\overline{\mathcal{T}}_{2j}^{n,i} = \phi_j \int_0^t \int_{[0,l]\times[0,\infty)} a\psi_j(\kappa_i,\epsilon_i) 1_{\{\kappa_i\}}(x) e^{-\lambda_j n(t-s)} \frac{1}{n} N_c^{n\varphi_n}(ds,dx,da),$$

$$\begin{aligned}\widetilde{\mathcal{T}}_{2j}^{n,i} &= \phi_j \int_0^t \int_{[0,l]\times[0,\infty)} a\psi_j(\kappa_i,\epsilon_i) 1_{\{\kappa_i\}}(x) e^{-\lambda_j n(t-s)} \frac{1}{n} n\varphi_n(s,x,a)\nu_T(ds,dx,da) \\ &= \phi_j \int_0^t \int_0^\infty a\psi_j(\kappa_i,\epsilon_i) e^{-\lambda_j n(t-s)} \varphi_n^i(s,a) f_i F_i(da) ds,\end{aligned}$$

here $N_c^{n\varphi_n} = N^{n\varphi_n} - n\varphi_n\nu_1$ is the compensated random measure for $N^{n\varphi_n}$ and $\varphi_n^i(s,a) = \varphi_n(s,\kappa_i,a)$.

Assumption 3.3.1 guarantees that the condition in Lemma 3.3.2 holds. Thus Lemma 3.3.2 gives us

$$\begin{aligned}||\sum_{j=1}^\infty \widetilde{\mathcal{T}}_{2j}^{n,i}(t)||_0^2 &= \sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) f_i^2 \left(\int_0^t \int_0^\infty a e^{-\lambda_j n(t-s)} \varphi_n^i(s,a) F_i(da) ds \right)^2 \\ &\le \sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) f_i^2 S^n(j)^2.\end{aligned}$$

Since for each $j \in \mathbb{N}$, $S^n(j)$ goes to zero a.s., as $n \to \infty$, and $S^n(j)$ is a.s. uniformly bounded in j and n, and $\sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) < \infty$, we have by dominated convergence theorem that $\sum_{j=1}^\infty \widetilde{\mathcal{T}}_{2j}^{n,i}$ converges a.s. in $C([0,1]:H)$ to 0.

Next consider $\overline{\mathcal{T}}_{2j}^{n,i}$, we will show that $\bar{\mathbb{E}}(\sup_{t\in[0,1]} ||\sum_{j=1}^\infty \overline{\mathcal{T}}_{2j}^{n,i}||_0^2) \to 0$. Define

$$M_t^{n,j} = \int_0^t \int_{[0,l]\times[0,\infty)} a 1_{\{\kappa_i\}}(x) e^{\lambda_j ns} N_c^{n\varphi_n}(ds,dx,da),$$

then $M_t^{n,j}$ is a martingale. Note that

$$\overline{\mathcal{T}}_{2j}^{n,i}(t) = \phi_j \psi_j(\kappa_i,\epsilon_i) \left(\frac{1}{n} e^{-\lambda_j nt} M_t^{n,j} \right).$$

Using integration by parts, we have

$$\frac{1}{n} e^{-\lambda_j nt} M_t^{n,j} = \frac{1}{n} \int_0^t e^{-\lambda_j ns} dM_s^{n,j} - \lambda_j \int_0^t e^{-\lambda_j ns} M_s^{n,j} ds. \tag{3.3.7}$$

The first term is a martingale and thus by Doob's maximal inequality for martingales,

$$
\begin{aligned}
\bar{\mathbb{E}}(\frac{1}{n}\sup_{t\in[0,1]} \quad & \int_0^t \quad e^{-\lambda_j ns} dM_s^{n,j})^2 \\
&\leq \frac{4}{n^2}\bar{\mathbb{E}}\left(\int_0^1 e^{-\lambda_j ns} dM_s^{n,j}\right)^2 \\
&= \frac{4}{n^2}\bar{\mathbb{E}}\left(\int_0^1 e^{-\lambda_j ns}\int_{[0,l]\times[0,\infty)} a 1_{\{\kappa_i\}}(x) e^{\lambda_j ns} N_c^{n\varphi_n}(ds,dx,da)\right)^2 \\
&= \frac{4}{n^2}\bar{\mathbb{E}}\left(\int_0^1\int_{[0,l]\times[0,\infty)} a^2 1_{\{\kappa_i\}}(x) n\varphi_n(s,x,a)\nu_1(ds,dx,da)\right) \\
&= \frac{4}{n} f_i \bar{\mathbb{E}}\left(\int_0^1\int_{[0,\infty)} a^2 \varphi_n^i(s,a) F_i(da) ds\right) \\
&\leq \frac{4}{n} f_i \left(\int_0^\infty e^{\delta a^2} F_i(da) + \bar{\mathbb{E}}\frac{1}{\delta}\int_0^1\int_0^\infty l(\varphi_n^i(s,a)) F_i(da) ds\right).
\end{aligned}
$$

The last inequality is once more a consequence of (3.3.5). Using Assumption 3.3.1 and the fact that $\varphi_n \in S^N$, we have that $\bar{\mathbb{E}}(\frac{1}{n}\sup_{t\in[0,1]}\int_0^t e^{-\lambda_j ns} dM_s^{n,j})^2$ goes to 0 as $n \to \infty$ and is uniformly bounded in j and n. Using the summability property $\sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) < \infty$, we now have that

$$
\sum_{j=1}^\infty \phi_j \psi_j(\kappa_i,\epsilon_i)\left(\frac{1}{n}\int_0^t e^{-\lambda_j ns} dM_s^{n,j}\right) \to 0, \tag{3.3.8}
$$

in probability in $D([0,1]:H)$, as $n \to \infty$.

Next consider the term $\lambda_j \int_0^t e^{-\lambda_j ns} M_s^{n,j} ds$, first we show $\bar{\mathbb{E}}(\lambda_j \int_0^t e^{-\lambda_j ns} M_s^{n,j} ds)^2 \to 0$ for all $t \in [0,1]$ and is bounded in j and n. Observe that if M_t is a square integrable martingale, then for $\alpha : [0,1] \to \mathbb{R}$ with suitable integrability properties,

$$
\bar{\mathbb{E}}\left(\int_0^t \alpha_s M_s ds\right)^2 = 2\int_0^t\int_0^s \alpha_s\alpha_u \bar{\mathbb{E}} M_s M_u du ds = 2\int_0^t\int_0^s \alpha_s\alpha_u \bar{\mathbb{E}}(M_u)^2 du ds. \tag{3.3.9}
$$

Using this observation,

$$
\begin{aligned}
&\bar{\mathbb{E}}\left(\lambda_j \int_0^t e^{-\lambda_j ns} M_s^{n,j} ds\right)^2 \\
&\quad = 2\lambda_j^2 \int_0^t \int_0^s e^{-\lambda_j ns} e^{-\lambda_j nu} \bar{\mathbb{E}}(M_u^{n,j})^2 duds \\
&\quad = 2\lambda_j^2 \int_0^t \int_0^s e^{-\lambda_j ns} e^{-\lambda_j nu} \bar{\mathbb{E}}\left(\int_0^u \int_0^\infty a^2 e^{2\lambda_j nv} n\varphi_n^i(v,a) f_i F_i(da) dv\right) duds \\
&\quad = 2nf_i\lambda_j^2 \bar{\mathbb{E}}\left(\int_0^t \int_0^\infty a^2 e^{2\lambda_j nv} \varphi_n^i(v,a) \int_v^t e^{-\lambda_j nu} \int_u^t e^{-\lambda_j ns} dsdudvF_i(da)\right) \\
&\quad = \frac{f_i}{n} \bar{\mathbb{E}}\left(\int_0^t \int_0^\infty (e^{-\lambda_j n(t-v)} - 1)^2 a^2 \varphi_n^i(v,a) F_i(da) dv\right) \\
&\quad \leq \frac{f_i}{n} \bar{\mathbb{E}}\left(\int_0^t \int_0^\infty a^2 \varphi_n^i(v,a) F_i(da) dv\right).
\end{aligned}
$$

Using similar argument as used in obtaining (3.3.8), we have $\bar{\mathbb{E}}\left(\lambda_j \int_0^t e^{-\lambda_j ns} M_s^{n,j} ds\right)^2$ goes to 0 as $n \to \infty$ and is bounded in j, n and t. This proves that

$$
Z^n(t) \equiv \sum_{j=1}^\infty \phi_j \psi_j(\kappa_i, \epsilon_i) \left(\lambda_j \int_0^t e^{-\lambda_j ns} M_s^{n,j} ds\right) \to 0, \tag{3.3.10}
$$

in probability, as $n \to \infty$ for each t. In order to prove uniform convergence, we use Aldous and Kurtz tightness criteria (see [53]). Given τ a stopping time such that $\tau \leq M$ a.s. for some constant $M \in (0,1]$, note that

$$
\begin{aligned}
&\bar{\mathbb{E}}||Z^n(\tau+\delta) - Z^n(\tau)||_0^2 \\
&\quad = \sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) \bar{\mathbb{E}}\left(\lambda_j \int_\tau^{\tau+\delta} e^{-\lambda_j ns} M_s^{n,j} ds\right)^2 \\
&\quad = \sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) \lambda_j^2 \bar{\mathbb{E}}\left(\int_0^1 1_{[\tau,\tau+\delta]}(s) e^{-\lambda_j ns} M_s^{n,j} ds\right)^2 \\
&\quad = 2\sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) \lambda_j^2 \bar{\mathbb{E}}\left(\int_0^1 1_{[\tau,\tau+\delta]}(s) e^{-\lambda_j ns} \int_0^s 1_{[\tau,\tau+\delta]}(u) e^{-\lambda_j nu} \bar{\mathbb{E}}[(M_u^{n,j})^2 | \mathscr{F}_\tau] duds\right) \\
&\quad \leq 2\sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) \lambda_j^2 \left(\int_0^1 e^{-\lambda_j ns} \int_0^s e^{-\lambda_j nu} \bar{\mathbb{E}}(M_u^{n,j})^2 duds\right) \\
&\quad \leq \frac{f_i}{n} \sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i) \bar{\mathbb{E}}\left(\int_0^1 \int_0^\infty a^2 \varphi_n^i(v,a) F_i(da) dv\right).
\end{aligned}
$$

Thus

$$\lim_{\delta \to 0} \overline{\lim_{n\to\infty}} \bar{\mathbb{E}} ||Z^n(\tau+\delta) - Z^n(\tau)||_0^2 = 0,$$

which along with the pointwise convergence in (3.3.10) yields that $Z^n \to 0$ in probability in $C([0,1]:H)$. Combining this with (3.3.7) and (3.3.8), we have $\sum_{j=1}^{\infty} \overline{\mathcal{T}}_{2j}^{n,i} \to 0$ in probability in $D([0,1]:H)$. Thus we have shown that $\mathcal{T}_2^n \to 0$ in probability in $D([0,1]:H)$, as $n \to \infty$.

We now show that $\mathcal{T}_1^n \Rightarrow \mathcal{G}^0\left(\nu_1^{\varphi}\right)$. For this, by a standard argument based on Skorokhod representation theorem, we can assume without loss of generality that $\varphi_n \to \varphi$ a.s., as $n \to \infty$.

Then, from Lemma 3.3.1, a.s.,

$$\int_0^t \int_0^\infty a\varphi_n^i(s,a)F_i(da)ds \to \int_0^t \int_0^\infty a\varphi^i(s,a)F_i(da)ds, \tag{3.3.11}$$

uniformly on $[0,1]$. Thus

$$\begin{aligned}\mathcal{T}_1^n(t) - \mathcal{G}^0\left(\nu_1^{\varphi}\right)(t) =& \phi_0 \int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_0]\frac{1}{n}N_c^{n\varphi_n}(ds,dx,da) \\ &+ \phi_0 \left(\int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_0]\varphi_n(s,x,a)\nu_1(ds,dx,da) \right. \\ &\left. - \int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_0]\varphi(s,x,a)\nu_1(ds,dx,da)\right).\end{aligned}$$

From (3.3.11), the second term goes to zero uniformly on $[0,1]$ a.s as $n \to \infty$. For the first term, notice that

$$\begin{aligned}\bar{\mathbb{E}} \left| \sup_{t\in[0,1]} \int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_0] \ \frac{1}{n}N_c^{n\varphi_n}(ds,dx,da)\right|^2 & \\ \le \frac{4}{n^2}\bar{\mathbb{E}} \int_{[0,1]\times[0,l]\times[0,\infty)} & (G(x,a)[\phi_0])^2 n\varphi_n\nu_1(ds,dx,da) \\ = \frac{4}{n} f_i \psi_0^2(\kappa_i,\epsilon_i)\bar{\mathbb{E}} & \int_0^1 \int_0^\infty a^2\varphi_n^i(s,a)F_i(da)ds.\end{aligned}$$

The expression on the right side converges to 0 as $n \to \infty$, since

$$\sup_{n\in\mathbb{N}} \mathbb{E} \int_0^1 \int_0^\infty a^2\varphi_n^i(s,a)F_i(da)ds < \infty,$$

from Assumption 3.3.1 (see the estimate in (3.3.8)). Thus $\mathcal{T}_1^n$ converges in $D([0,1]:H)$ in probability to $\mathcal{G}^0(\nu_1^{\varphi})$. The result follows. $\square$

Finally we can complete the proof of Theorem 3.3.2.

Proof of Theorem 3.3.2. From Proposition 3.3.1 and 3.3.2, we see that Condition 2.3.2 is satisfied. Theorem now is an immediate consequence of Theorem 2.3.2. $\square$

3.4 The case $\alpha > 0$.

In this section, we will assume that $\alpha > 0$. This condition says that the chemical amounts dissipate at a strictly positive rate. The following theorem from [52] shows that, in this case, u_t converges weakly to a random field.

Theorem 3.4.1. *If $u_0 \in \Phi_0$, then u_t converges weakly in Φ_0 to a random field u_∞. For $x \in [0,l]$, let $u(x) = \mathbb{E}u_\infty(x)$. Then $u(x)$ is the solution of the following differential equation*

$$D\frac{d^2u(x)}{dx^2} - V\frac{du(x)}{dx} - \alpha u(x) + Q(x) = 0, \tag{3.4.1}$$

where

$$Q(x) = \sum_{i=1}^{r} \frac{a_i f_i}{2\epsilon_i} 1_{(\kappa_i-\epsilon_i,\kappa_i+\epsilon_i)}(x).$$

Remark 3.4.1. By a solution of (3.4.1) we mean a $u \in \Phi_0$ such that $\forall \phi \in \Phi$

$$\langle u, L\phi\rangle_0 - \alpha\langle u, \phi\rangle_0 + \sum_{i=1}^{r} \frac{a_i f_i}{2\epsilon_i} \int_{\kappa_i-\epsilon_i}^{\kappa_i+\epsilon_i} \phi(y)\rho(y)dy = 0.$$

The above result in particular shows that as $t \to \infty$, $\frac{u_t}{t} \to 0$ in Φ_0, in probability and thus in order to obtain a nontrivial time asymptotic result it is more natural to consider time averages of the form $\frac{1}{T}\int_0^T u_s ds$. The following result establishes a law of large numbers for such time averages.

Theorem 3.4.2. *Let $u_0 \in \Phi_0$. Then, as $T \to \infty$, $\frac{1}{T}\int_0^T u_t dt$ converges to u in Φ_0, in probability.*

Proof. From (3.2.6) and recalling that $\{\phi_j\}$ are eigenfunctions of L, it follows that

$$\begin{aligned} u_t[\phi_j] =& e^{-(\alpha+\lambda_j)t} u_0[\phi_j] + \sum_{i=1}^r f_i a_i \psi_j(\kappa_i, \epsilon_i) e^{-(\alpha+\lambda_j)t} \int_0^t e^{(\alpha+\lambda_j)s} ds \\ &+ e^{-(\alpha+\lambda_j)t} \int_{[0,t]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)s} G(x,a)[\phi_j] N_c(ds,dx,da). \end{aligned}$$

Thus

$$\begin{aligned} \frac{1}{T}\int_0^T u_t dt =& \frac{1}{T}\int_0^T \left(\sum_{j=0}^\infty e^{-(\alpha+\lambda_j)t} u_0[\phi_j]\phi_j\right) dt \\ &+ \sum_{j=0}^\infty \phi_j \sum_{i=1}^r f_i a_i \psi_j(\kappa_i,\epsilon_i) \frac{1}{T}\int_0^T \frac{1-e^{-(\alpha+\lambda_j)t}}{\alpha+\lambda_j} dt \\ &+ \sum_{j=0}^\infty \phi_j \frac{1}{T}\int_0^T e^{-(\alpha+\lambda_j)t} \int_{[0,t]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)s} G(x,a)[\phi_j] N_c(ds,dx,da) dt \\ =& \mathcal{T}_1 + \mathcal{T}_2 + \mathcal{T}_3. \end{aligned} \tag{3.4.2}$$

For the first term note that

$$||\mathcal{T}_1||_0 = ||\frac{1}{T}\int_0^T e^{-\alpha t}(T_t u_0) dt||_0 \le \frac{1}{T}\int_0^T e^{-\alpha t} ||(T_t u_0)||_0 dt \le ||u_0||_0 \frac{1-e^{-\alpha T}}{T\alpha} \to 0 \tag{3.4.3}$$

as $T \to \infty$. Therefore, $\mathcal{T}_1$ goes to zero.

We next show that $\mathcal{T}_3$ goes to zero in probability. Write $\mathcal{T}_3$ as

$$\mathcal{T}_3 = \sum_{j=0}^\infty \left(\frac{1}{T}\int_0^T \alpha_t^j M_t^j dt\right) \phi_j,$$

where

$$\begin{aligned} \alpha_t^j &= e^{-(\alpha+\lambda_j)t} \\ M_t^j &= \int_{[0,t]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)s} G(x,a)[\phi_j] N_c(ds,dx,da). \end{aligned}$$

Noting that M_t^j is a martingale for each j, we have from (3.3.9) that,

$$\mathbb{E}\left(\frac{1}{T}\int_0^T \alpha_t^j M_t^j dt\right)^2 = \frac{2}{T^2}\int_0^T\int_0^t e^{-(\alpha+\lambda_j)t}e^{-(\alpha+\lambda_j)s}\mathbb{E}(M_s^j)^2 dsdt.$$

Also,

$$\begin{aligned}\mathbb{E}(M_s^j)^2 &= \int_{[0,s]\times[0,l]\times[0,\infty)} e^{2(\alpha+\lambda_j)u}G^2(x,a)[\phi_j]\nu(du,dx,da)\\ &= \sum_{i=1}^r \psi_j^2(\kappa_i,\epsilon_i)f_i\int_{[0,s]\times[0,\infty)} e^{2(\alpha+\lambda_j)u}a^2F_i(da)du\\ &= \sum_{i=1}^r \psi_j^2(\kappa_i,\epsilon_i)f_ib_i\frac{e^{2(\alpha+\lambda_j)s}-1}{2(\alpha+\lambda_j)},\end{aligned}$$

where recall that $b_i = \int_{[0,\infty)} a^2F_i(da)$.

Thus

$$\begin{aligned}&\mathbb{E}\left(\frac{1}{T}\int_0^T \alpha_t^j M_t^j dt\right)^2\\ &= \frac{2}{T^2}\sum_{i=1}^r \psi_j^2(\kappa_i,\epsilon_i)f_ib_i\int_0^T\int_0^t e^{-(\alpha+\lambda_j)t}e^{-(\alpha+\lambda_j)s}\frac{e^{2(\alpha+\lambda_j)s}-1}{2(\alpha+\lambda_j)}dsdt\\ &= \frac{2}{T^2}\left(\frac{T}{2(\alpha+\lambda_j)^2}-\frac{e^{-2(\alpha+\lambda_j)T}-1}{4(\alpha+\lambda_j)^3}+\frac{e^{-(\alpha+\lambda_j)T}-1}{(\alpha+\lambda_j)^3}\right)\sum_{i=1}^r \psi_j^2(\kappa_i,\epsilon_i)f_ib_i\\ &\le \frac{2}{T^2}\left(\frac{T}{2\alpha^2}+\frac{1}{4\alpha^3}+\frac{1}{\alpha^3}\right)\sum_{i=1}^r \psi_j^2(\kappa_i,\epsilon_i)f_ib_i \to 0,\end{aligned}$$

as $T\to\infty$. And since $\sum_{j=1}^\infty \psi_j^2(\kappa_i,\epsilon_i)<\infty$, $\forall i=1,...,r$, we have

$$\mathbb{E}||\mathcal{T}_3||_0^2 = \sum_{j=0}^\infty \mathbb{E}\left(\frac{1}{T}\int_0^T \alpha_t^j M_t^j dt\right)^2 \to 0. \tag{3.4.4}$$

Finally we argue that $\mathcal{T}_2 \to u$ in Φ_0. Recall that

$$\mathcal{T}_2 = \sum_{j=0}^\infty \phi_j \sum_{i=1}^r f_ia_i\psi_j(\kappa_i,\epsilon_i)\frac{1}{T}\int_0^T \frac{1-e^{-(\alpha+\lambda_j)t}}{\alpha+\lambda_j}dt.$$

Also, from Remark 3.4.1

$$u = \sum_{j=0}^\infty \phi_j \sum_{i=1}^r f_ia_i\psi_j(\kappa_i,\epsilon_i)\frac{1}{\alpha+\lambda_j}.$$

Thus

$$
\begin{aligned}
||\mathcal{T}_2 - u||_0^2 &= \sum_{j=0}^{\infty} \left(\sum_{i=1}^{r} f_i a_i \psi_j(\kappa_i, \epsilon_i) \frac{1}{T} \int_0^T \frac{e^{-(\alpha+\lambda_j)t}}{\alpha+\lambda_j} dt \right)^2 \\
&\leq r \sum_{j=0}^{\infty} \left(\sum_{i=1}^{r} f_i^2 a_i^2 \psi_j^2(\kappa_i, \epsilon_i) \right) \left(\frac{1}{T} \int_0^T \frac{e^{-(\alpha+\lambda_j)t}}{\alpha+\lambda_j} dt \right)^2 \\
&= r \sum_{j=0}^{\infty} \left(\sum_{i=1}^{r} f_i^2 a_i^2 \psi_j^2(\kappa_i, \epsilon_i) \right) \left(\frac{1}{T} \frac{1 - e^{-(\alpha+\lambda_j)T}}{(\alpha+\lambda_j)^2} \right)^2 \\
&\leq \frac{r}{T^2 \alpha^4} \sum_{j=0}^{\infty} \left(\sum_{i=1}^{r} f_i^2 a_i^2 \psi_j^2(\kappa_i, \epsilon_i) \right) \to 0,
\end{aligned}
\tag{3.4.5}
$$

as $T \to \infty$, using the fact that $\sum_{j=1}^{\infty} \psi_j^2(\kappa_i, \epsilon_i) < \infty$. The result follows upon using (3.4.3),(3.4.4) and (3.4.5) in (3.4.2). $\square$

Next we will study the large deviation behavior of $\frac{1}{T}\int_0^T u_t dt$ as T becomes large. Equivalently, one can consider the family $\{\frac{1}{n}\int_0^{nt} u_s ds : t \in [0,1]\}$ as n becomes large. We begin by observing that for $n \in \mathbb{N}$, $\{\tilde{u}_t^n\}_{t\in[0,1]}$ given as the pathwise solution of the integral equation

$$
\begin{aligned}
\tilde{u}_t^n = &\int_0^t e^{-\alpha n s} T_{ns} u_0 ds \\
&+ n \sum_{j=0}^{\infty} \phi_j \int_0^t e^{-(\alpha+\lambda_j)ns} \int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu} G(x,a)[\phi_j] \eta_n(du, dx, da)
\end{aligned}
$$

on the probability space $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}), \bar{\mathscr{F}}_t)$, where $\eta_n = \frac{1}{n}N^n$, has the same distribution as $\{\frac{1}{n}\int_0^{nt} u_s ds\}_{t\in[0,1]}$. Here, as in Section 3.3, $\mathbb{X} = [0,l] \times [0,\infty)$, and $\bar{\mathbb{M}}$, $\mathcal{B}(\bar{\mathbb{M}})$, $\bar{\mathscr{F}}_t$ and $\frac{1}{n}N^n$ are defined in the same way as in Section 2.2.2. We will now study the large deviation behavior of the sequence of $\mathbb{U} = D([0,1] : H)$ valued random variables $\{\tilde{u}^n\}_{n\in\mathbb{N}}$.

Note that one can represent, for each $n \in \mathbb{N}$, $\tilde{u}^n = \mathcal{G}^n(\frac{1}{n}N^n)$ for some measurable map $\mathcal{G}^n : \mathbb{M} \to \mathbb{U}$. Define the map $\mathcal{G}^0 : \mathbb{M} \to \mathbb{U}$ as follows. For $\nu \in \mathbb{M}$, let

$$
\mathcal{G}^0(\nu)(t) = \sum_{j=0}^{\infty} \phi_j \frac{1}{\alpha+\lambda_j} \int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_j] \nu(ds, dx, da), \quad t \in [0,1];
$$

if $\sum_{j=0}^{\infty} \frac{1}{(\alpha+\lambda_j)^2} \left| \int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_j]\nu(ds,dx,da) \right|^2 < \infty$. If the latter integral is infinite, we set $\mathcal{G}^0(\nu) = 0$ (i.e. the zero trajectory in $D([0,1] : H)$). Define I through (2.3.4).

The following is the main result of this section.

Theorem 3.4.3. *Under Assumption 3.3.1, I is a rate function on $\mathbb{U}$ and the family $\{\tilde{u}^n\}_{n\in\mathbb{N}}$ satisfies a large deviation principle, as $n \to \infty$, on $D([0,1] : H)$, with rate function I.*

As in Section 3.3, the key step in the proof of Theorem 3.4.3 is once more the verification of Condition 2.3.2. We first consider Part 1 of the condition.

Part 1 of Condition 2.3.2:

Let $g_n, g \in S^N$ be such that $g_n \to g$. We will like to show that $\mathcal{G}^0(\nu_1^{g_n}) \to \mathcal{G}^0(\nu_1^g)$. For $t \in [0,1]$ (integrability of the right side below is readily verified),

$$\begin{aligned}\mathcal{G}^0(\nu_1^{g_n})(t) &= \sum_{j=0}^{\infty} \phi_j \frac{1}{\alpha+\lambda_j} \int_{[0,t]\times[0,l]\times[0,\infty)} G(x,a)[\phi_j] g_n(s,x,a)\nu(ds,dx,da) \\ &= \sum_{j=0}^{\infty} \phi_j \frac{1}{\alpha+\lambda_j} \sum_{i=1}^{r} f_i \psi_j(\kappa_i,\epsilon_i) \int_{[0,t]\times[0,\infty)} a g_n^i(s,a) F_i(da) ds,\end{aligned}$$

where $g_n^i(s,a) = g_n(s,\kappa_i,a)$. Thus

$$\begin{aligned}&||\mathcal{G}^0(\nu_1^{g_n})(t) - \mathcal{G}^0(\nu_1^g)(t)||_0^2 \\ &= \sum_{j=0}^{\infty} \frac{1}{(\alpha+\lambda_j)^2} \left(\sum_{i=1}^{r} f_i \psi_j(\kappa_i,\epsilon_i) \int_{[0,t]\times[0,\infty)} a(g_n^i - g^i) F_i(da) ds \right)^2 \\ &\leq \sum_{j=0}^{\infty} \frac{1}{(\alpha+\lambda_j)^2} \left(\sum_{i=1}^{r} f_i^2 \psi_j^2(\kappa_i,\epsilon_i) \right) \left(\sum_{i=1}^{r} \left(\int_{[0,t]\times[0,\infty)} a(g_n^i - g^i) F_i(da) ds \right)^2 \right) \\ &\to 0,\end{aligned}$$

uniformly in $t \in [0,1]$, since from Lemma 3.3.1, $\int_{[0,t]\times[0,\infty)} a(g_n^i - g^i) F_i(da) ds$ goes to zero uniformly for every $i \in \{1, ..., r\}$, and $\sum_{j=1}^{\infty} \psi_j^2(\kappa_i,\epsilon_i) < \infty$. Thus we have shown the following result.

Proposition 3.4.1. *Under Assumption 3.3.1, for every $N \in \mathbb{N}$, and $g_n, g \in S^N$,$n \geq 1$, such that $g_n \to g$, we have*

$$\mathcal{G}^0\left(\nu_T^{g_n}\right) \to \mathcal{G}^0\left(\nu_T^{g}\right)$$

in $\mathbb{U}$.

Next, we proceed to verify Part 2 of Condition 2.3.2.

Part 2 of Condition 2.3.2:

Let $\varphi_n, \varphi \in \mathcal{U}^N$. Then, for $t \in [0,1]$,

$$\begin{aligned}\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right)(t) &= \int_0^t e^{-\alpha ns}T_{ns}u_0 ds \\ &+ n\sum_{j=0}^{\infty}\phi_j\int_0^t e^{-(\alpha+\lambda_j)ns}\int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu}G(x,a)[\phi_j]\frac{1}{n}N^{n\varphi_n}(du,dx,da),\end{aligned}$$

and

$$\mathcal{G}^0\left(\nu_T^{\varphi}\right)(t) = \sum_{j=0}^{\infty}\phi_j\frac{1}{\alpha+\lambda_j}\sum_{i=1}^{r}f_i\psi_j(\kappa_i,\epsilon_i)\int_{[0,t]\times[0,\infty)} a\varphi^i(s,a)F_i(da)ds.$$

Proposition 3.4.2. *Let $\varphi_n, \varphi \in \mathcal{U}^N$ be such that $\varphi_n \Rightarrow \varphi$, as $n \to \infty$. Then under Assumption 3.3.1, we have*

$$\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right) \Rightarrow \mathcal{G}^0\left(\nu_T^{\varphi}\right).$$

Proof. Write

$$\mathcal{G}^n\left(\frac{1}{n}N^{n\varphi_n}\right)(t) = \mathcal{T}_0^n(t) + \mathcal{T}_1^n(t) + \mathcal{T}_2^n(t),$$

where, for $t \in [0,1]$,

$$\mathcal{T}_0^n(t) = \int_0^t e^{-\alpha ns} T_{ns} u_0 ds,$$
$$\mathcal{T}_1^n(t) = \sum_{j=0}^{\infty} \phi_j \int_0^t e^{-(\alpha+\lambda_j)ns} \int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu} G(x,a)[\phi_j] N_c^{n\varphi_n}(du,dx,da),$$
$$\mathcal{T}_2^n(t) = \sum_{j=0}^{\infty} \phi_j \int_0^t e^{-(\alpha+\lambda_j)ns} \int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu} G(x,a)[\phi_j] n\varphi_n(s,x,a)\nu(du,dx,da).$$

For the first term note that, for $t \in [0,1]$,

$$\begin{aligned}||\mathcal{T}_0^n(t)||_0 = ||\int_0^t e^{-\alpha ns} T_{ns} u_0 ds||_0 &\le \int_0^t e^{-\alpha ns} ||T_{ns}u_0||_0 ds \\ &\le ||u_0||_0 \int_0^t e^{-\alpha ns} ds = ||u_0||_0 \frac{1-e^{-\alpha nt}}{\alpha n} \le \frac{||u_0||_0}{\alpha n} \to 0,\end{aligned}$$

as $n \to \infty$. Thus $\mathcal{T}_0^n(t) \to 0$ uniformly in $t \in [0,1]$.

Next, write $\mathcal{T}_1^n(t)$ as $\mathcal{T}_1^n(t) = \sum_{j=0}^{\infty} \phi_j \sum_{i=1}^{r} \mathcal{T}_{1j}^{n,i}(t)$, where

$$\begin{aligned}\mathcal{T}_{1j}^{n,i}(t) &= \int_0^t e^{-(\alpha+\lambda_j)ns} \int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu} 1_{\kappa_i}(x) G(x,a)[\phi_j] N_c^{n\varphi_n}(du,dx,da) \\ &= \psi_j(\kappa_i,\epsilon_i) \int_0^t e^{-(\alpha+\lambda_j)ns} \int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu} 1_{\kappa_i}(x) a N_c^{n\varphi_n}(du,dx,da).\end{aligned}$$

Using similar argument as in the proof of Theorem 3.4.2, we get

$$\begin{aligned}&\bar{\mathbb{E}}(\mathcal{T}_{1j}^{n,i}(t))^2 \\ =&\bar{\mathbb{E}}\left(\int_0^t e^{-(\alpha+\lambda_j)ns} \int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu} 1_{\kappa_i}(x) G(x,a)[\phi_j] N_c^{n\varphi_n}(du,dx,da)\right)^2 \\ =&2\bar{\mathbb{E}}\left(\int_0^t \int_0^s e^{-(\alpha+\lambda_j)ns} e^{-(\alpha+\lambda_j)nu} \right. \\ &\left. \int_{[0,u]\times[0,l]\times[0,\infty)} e^{2(\alpha+\lambda_j)nv} 1_{\kappa_i}(x) G^2(x,a)[\phi_j] n\varphi_n(s,x,a)\nu(dv,dx,da)duds\right) \\ =&2nf_i\psi_j^2(\kappa_i,\epsilon_i)\bar{\mathbb{E}}\left(\int_0^t \int_0^s e^{-(\alpha+\lambda_j)ns} e^{-(\alpha+\lambda_j)nu} \right. \\ &\left. \int_{[0,u]\times[0,\infty)} e^{2(\alpha+\lambda_j)nv} a^2 \varphi_n^i(v,a) F_i(da) dvduds\right).\end{aligned}$$

Changing the order of integration, after some calculations, we have

$$
\begin{aligned}
&\bar{\mathbb{E}}(\mathcal{T}_{1j}^{n,i}(t))^2\\
=&\frac{f_i\psi_j^2(\kappa_i,\epsilon_i)}{(\alpha+\lambda_j)^2 n}\bar{\mathbb{E}}\left(\int_{[0,t]\times[0,\infty)}(1+e^{-2(\alpha+\lambda_j)n(t-v)}-2e^{-(\alpha+\lambda_j)n(t-v)})a^2\varphi_n^i(v,a)F_i(da)dv\right)\\
\leq&\frac{f_i\psi_j^2(\kappa_i,\epsilon_i)}{(\alpha+\lambda_j)^2 n}\bar{\mathbb{E}}\left(\int_{[0,t]\times[0,\infty)}a^2\varphi_n^i(v,a)F_i(da)dv\right).
\end{aligned}
$$

Also note that

$$
\begin{aligned}
\int_{[0,t]\times[0,\infty)}a^2\varphi_n^i(v,a)F_i(da)dv &\leq \int_{[0,t]\times[0,\infty)}e^{\delta a^2}F_i(da)dv+\frac{1}{\delta}\int_{[0,t]\times[0,\infty)}l(\varphi_n^i)F_i(da)dv\\
&\leq \int_{[0,\infty)}e^{\delta a^2}F_i(da)+\frac{N}{\delta f_i}\quad a.s.
\end{aligned}
$$

Thus $\bar{\mathbb{E}}(\mathcal{T}_{1j}^{n,i}(t))^2 \to 0$ as $n\to\infty$. Combining this with the observation that $\sum_{j=0}^{\infty}\psi_j^2(\kappa_i,\epsilon_i)<\infty$, we now have that $\bar{\mathbb{E}}||\mathcal{T}_1^n(t)||^2\to 0$ for all $t\in[0,1]$.

In order to show uniform convergence of $\mathcal{T}_1^n$ to zero, we will use Aldous' tightness criteria (see [6]). Given τ a stopping time such that $\tau\leq M$ a.s. for a constant M, note that

$$
\begin{aligned}
&\bar{\mathbb{E}}(\mathcal{T}_{1j}^{n,i}(\tau+\delta)-\mathcal{T}_{1j}^{n,i}(\tau))^2\\
&=\psi_j^2(\kappa_i,\epsilon_i)\bar{\mathbb{E}}\left(\int_{\tau}^{\tau+\delta}e^{-(\alpha+\lambda_j)ns}\int_{[0,s]\times[0,l]\times[0,\infty)}e^{(\alpha+\lambda_j)nu}1_{\kappa_i}(x)aN_c^{n\varphi_n}(du,dx,da)ds\right)^2\\
&\leq \psi_j^2(\kappa_i,\epsilon_i)\bar{\mathbb{E}}\left[\left(\int_0^1 1_{[\tau,\tau+\delta]}(s)ds\right)\right.\\
&\qquad\left.\times\left(\int_0^1 e^{-2(\alpha+\lambda_j)ns}\left(\int_{[0,s]\times[0,l]\times[0,\infty)}e^{(\alpha+\lambda_j)nu}1_{\kappa_i}(x)aN_c^{n\varphi_n}(du,dx,da)\right)^2ds\right)\right]\\
&\leq \psi_j^2(\kappa_i,\epsilon_i)\delta\int_0^1 e^{-2(\alpha+\lambda_j)ns}\bar{\mathbb{E}}\left(\int_{[0,s]\times[0,l]\times[0,\infty)}e^{(\alpha+\lambda_j)nu}1_{\kappa_i}(x)aN_c^{n\varphi_n}(du,dx,da)\right)^2ds\\
&=\psi_j^2(\kappa_i,\epsilon_i)\delta\int_0^1 e^{-2(\alpha+\lambda_j)ns}\bar{\mathbb{E}}\left(\int_{[0,s]\times[0,\infty)}e^{2(\alpha+\lambda_j)nu}a^2n\varphi_n^i(u,a)f_iF_i(da)du\right)ds\\
&\leq \frac{\psi_j^2(\kappa_i,\epsilon_i)f_i\delta}{2(\alpha+\lambda_j)}\bar{\mathbb{E}}\left(\int_{[0,1]\times[0,\infty)}a^2\varphi_n^i(u,a)F_i(da)du\right).
\end{aligned}
$$

Using the summability of $\psi_j^2(\kappa_i,\epsilon_i)$ again, we now see that

$$
\bar{\mathbb{E}}||\mathcal{T}_1^n(\tau+\delta)-\mathcal{T}_1^n(\tau)||_0^2\leq c_0\delta,
$$

where c_0 depends only on M. Tightness of $\{\mathcal{T}_1^n\}_{n\in\mathbb{N}}$ in $D([0,1]:\Phi_0)$ follows. Combing with pointwise convergence to 0, we now have that $\mathcal{T}_1^n$ converges in probability to 0 in $D([0,1]:\Phi_0)$.

Finally, consider $\mathcal{T}_2^n$:

$$
\begin{aligned}
&\mathcal{T}_2^n(t)\\
&= n\sum_{j=0}^{\infty}\phi_j\int_0^t e^{-(\alpha+\lambda_j)ns}\int_{[0,s]\times[0,l]\times[0,\infty)} e^{(\alpha+\lambda_j)nu}G(x,a)[\phi_j]\varphi_n(s,x,a)\nu(du,dx,da)\\
&= n\sum_{j=0}^{\infty}\phi_j\int_0^t e^{-(\alpha+\lambda_j)ns}\sum_{i=1}^{r}\int_{[0,s]\times[0,\infty)} e^{(\alpha+\lambda_j)nu}a\psi_j(\kappa_i,\epsilon_i)\varphi_n^i(u,a)f_iF_i(da)duds\\
&= n\sum_{j=0}^{\infty}\phi_j\sum_{i=1}^{r}\psi_j(\kappa_i,\epsilon_i)f_i\int_{[0,t]\times[0,\infty)} e^{(\alpha+\lambda_j)nu}\int_u^t e^{-(\alpha+\lambda_j)ns}a\varphi_n^i(u,a)F_i(da)dsdu\\
&= \sum_{j=0}^{\infty}\frac{1}{\alpha+\lambda_j}\phi_j\sum_{i=1}^{r}\psi_j(\kappa_i,\epsilon_i)f_i\int_{[0,t]\times[0,\infty)}(1-e^{-(\alpha+\lambda_j)n(t-u)})a\varphi_n^i(u,a)F_i(da)du.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&||\mathcal{T}_2^n(t)-\mathcal{G}^0\left(\nu_1^{\varphi}\right)(t)||_0^2\\
&= \sum_{j=0}^{\infty}\frac{1}{(\alpha+\lambda_j)^2}\left(\sum_{i=1}^{r}\psi_j(\kappa_i,\epsilon_i)f_i\int_{[0,t]\times[0,\infty)}(1-e^{-(\alpha+\lambda_j)n(t-u)})a\varphi_n^i(u,a)F_i(da)du\right)^2\\
&\le \sum_{j=0}^{\infty}\frac{1}{(\alpha+\lambda_j)^2}\left(\sum_{i=1}^{r}\psi_j^2(\kappa_i,\epsilon_i)f_i^2\right)\\
&\qquad\times\left(\sum_{i=1}^{r}\left(\int_{[0,t]\times[0,\infty)}(1-e^{-(\alpha+\lambda_j)n(t-u)})a\varphi_n^i(u,a)F_i(da)du\right)^2\right).
\end{aligned}
$$

Once again, applying the Skorokhod representation theorem, we can assume without loss of generality that $\varphi_n\to\varphi$ a.s., as $n\to\infty$. Then, using Lemma 3.3.1 and Lemma 3.3.2 (with λ_j replaced by $\alpha+\lambda_j$ therein),

$$
\int_0^t\int_0^{\infty} a(1-e^{-(\alpha+\lambda_j)n(t-u)})\varphi_n^i(u,a)F_i(da)du\to\int_0^t\int_0^{\infty}a\varphi^i(u,a)F_i(da)du \quad (3.4.6)
$$

uniformly for $t\in[0,1]$, a.s., as $n\to\infty$.

Thus we have $\mathcal{T}_2^n\Rightarrow\mathcal{G}^0\left(\nu_1^{\varphi}\right)$ in $D([0,1]:H)$. The proposition is proved. $\square$

Finally we can complete the proof of Theorem 3.4.3.

Proof of Theorem 3.4.3. From Proposition 3.4.1 and 3.4.2, we see that Condition 2.3.2 is satisfied. Theorem now is an immediate consequence of Theorem 2.3.2. $\square$

Chapter 4

Large Deviations for Stochastic Partial Differential Equations Driven by a Poisson Random Measure

4.1 Introduction.

Stochastic partial differential equations driven by Poisson random measures arise in many different fields. For example, they have been used to develop models for neuronal activity that account for synaptic impulses occurring randomly, both in time and at different locations of a spatially extended neuron. Other applications arise in chemical reaction-diffusion systems and stochastic turbulence models. The starting point in all these application areas are deterministic partial differential equations (PDE) that capture the underlying physics. One then develops a stochastic evolution model driven by a suitable Poisson noise process to take into account random inputs or effects to the nominal deterministic dynamics. In typical settings the solutions of these stochastic evolution equations are not smooth. In fact in many applications of interest they are not even random fields (that is, function valued), and therefore an appropriate framework is given through the theory of generalized functions. A systematic theory of existence and uniqueness of solutions (both weak and pathwise) for such stochastic partial differential equations (SPDE) driven by Poisson random measures has been developed in [52]. Our objective in this chapter is to study some large deviation problems associated with such stochastic systems.

Large deviation properties of SPDE driven by infinite dimensional Brownian motions (e.g. Brownian sheets) have been extensively studied. In a typical such setting one considers a small parameter multiplying the noise term and is interested in asymptotic probabilities of non-nominal behavior as the parameter approaches zero. This is the classical Freidlin-Wentzell problem that has been studied in numerous papers (see the references in [16]). Earlier works on this family of problems were based on ideas of [2] and relied on discretizations and other approximations combined with 'super-exponential closeness' probability estimates. For many models of interest, particularly those arising from fluid dynamics and turbulence, developing the required exponential probability estimates is a daunting task and consequently simpler alternative methods are of interest. In recent years an approach based on certain variational representation formulas for moments of nonnegative functionals of Brownian motions [16] has been increasingly used for the study of the small noise large deviation problem for Brownian motion driven infinite dimensional systems [5, 16, 17, 23, 30, 31, 61, 63, 68, 69, 71, 74, 81, 86, 89, 90]. The main appealing feature of this approach is that it completely bypasses approximation/discretization arguments and exponential probability estimates, and in their place essentially requires a basic qualitative understanding of existence, uniqueness and stability (under 'bounded' perturbations) of certain controlled analogues of the underlying stochastic dynamical system of interest.

Large deviation results for finite dimensional stochastic differential equations with a Poisson noise term has been studied by several authors [83, 59, 33, 29]. For infinite dimensional models with jumps, very little is available. One exception is the paper [70] that obtains large deviation results for an Ornstein-Uhlenbeck type process driven by an infinite dimensional Lévy noise. One reason there is relatively little work in the Poisson noise setting is that approximation arguments that one uses for Brownian noise models become much more onerous in the Poisson setting, and for general infinite dimensional models the approach of [2] becomes intractable.

With the expectation that it would prove useful for the study of large deviations for SPDEs driven by Poisson Random Measures (PRM), the paper [18] developed a variational representation, for moments of non negative functionals of PRMs, which is analogous to the representation given in [14, 16] for the Brownian motion case. The paper [18] also obtained large deviation results for a basic model of a finite dimensional jump-diffusion to illustrate the applicability of this variational representation for the study of large deviation problems for models with jumps. However the feasibility of this approach for the study of complex infinite dimensional stochastic dynamical systems driven by Poisson random measures has not been addressed to date.

The goal of this chapter is to demonstrate that the approach based on variational representations that has been very successful for obtaining large deviation results for system driven by Brownian noises works equally well for SPDE models driven by PRMs. As in the Brownian case we study the small noise problem, which in the Poisson setting means that the jump intensity is $O(\epsilon^{-1})$ and jump sizes are $O(\epsilon)$, where ϵ is a small parameter. We consider a rather general family of models of the form

$$X_t^\epsilon = X_0^\epsilon + \int_0^t A(s, X_s^\epsilon)ds + \epsilon \int_0^t \int_{\mathbb{X}} G(s, X_{s-}^\epsilon, v)\tilde{N}^{\epsilon^{-1}}(dsdv), \tag{4.1.1}$$

where $N^{\epsilon^{-1}}$ is a Poisson random measure on $[0,T] \times \mathbb{X}$ with a σ-finite mean measure $\epsilon^{-1}\lambda_T \otimes \nu$, λ_T is the Lebesgue measure on $[0,T]$ and $\tilde{N}^{\epsilon^{-1}}([0,t] \times B) = N^{\epsilon^{-1}}([0,t] \times B) - \epsilon^{-1}t\nu(B)$, $\forall B \in \mathcal{B}(\mathbb{X})$ with $\nu(B) < \infty$, is the compensated Poisson random measure.

As noted previously, a key issue with a Poisson noise model is the selection of an appropriate state space, since it is natural and often convenient for there to be little spatial regularity. However, many of these foundational issues have been satisfactorily resolved in [52], where pathwise existence and uniqueness of SPDE of the form (4.1.1) are treated under rather general conditions. In the framework of [52] solutions lie in the space of RCLL trajectories that take values in the dual of a suitable nuclear space. This framework covers many specific application settings that have been studied in the literature (e.g., spatially extended neuron models, chemical reaction-diffusion systems, etc.). In a parallel with the case of Brownian noise, one finds that the estimates needed for establishing the well-posedness of the equation are precisely the ones that are key for the proof of the large deviation result as well.

This Chapter is organized as follows. We begin in Section 4.2 with some background results. A strengthening of the general large deviation result established in [18] is presented. Also summarized are basic existence and uniqueness results from [52] for SPDEs with solutions in the duals of Countably Hilbertian Nuclear Spaces (CHNS). In Section 4.3 we study the small noise problem and state verifiable conditions on the model data in (4.1.1) under which a large deviations principle holds. Finally, the Appendix collects some auxiliary results.

4.2 Preliminaries.

4.2.1 A General Large Deviation Result.

Recall Condition 2.3.2 in Chapter 2. The first condition requires continuity in the control for deterministic controlled systems. The second condition is a law of large numbers result for small noise controlled stochastic systems. In both cases we are allowed to assume the controls take values in a compact set.When applied to the SDE (4.1.1), $\mathcal{G}^\epsilon$ will be the mapping that takes the PRM into X^ϵ.

Recall the rate function defined in (2.3.4). We will state it here again for easy reference. For $\phi \in \mathbb{U}$, define $\mathbb{S}_\phi = \{g \in \mathbb{S} : \phi = \mathcal{G}^0(\nu_T^g)\}$. Let $I : \mathbb{U} \to [0, \infty]$ be defined by

$$I(\phi) = \inf_{g \in \mathbb{S}_\phi} \{L_T(g)\}, \quad \phi \in \mathbb{U}. \tag{4.2.1}$$

By convention, $I(\phi) = \infty$ if $\mathbb{S}_\phi = \varnothing$.

For applications, the following strengthened form of Theorem 2.3.2 is useful. The proof follows by straightforward modifications; for completeness we include a sketch in the appendix of this chapter.

Let $\{K_n \subset \mathbb{X}, n = 1, 2, \ldots\}$ be an increasing sequence of compact sets such that $\cup_{n=1}^\infty K_n = \mathbb{X}$. For each n let

$$\begin{aligned} \bar{\mathcal{A}}_{b,n} \doteq \Big\{ & \varphi \in \bar{\mathcal{A}} : \text{ for all } (t,\omega) \in [0,T] \times \bar{\mathbb{M}},\ n \geq \varphi(t,x,\omega) \geq 1/n \text{ if } x \in K_n \\ & \text{and } \varphi(t,x,\omega) = 1 \text{ if } x \in K_n^c \Big\}, \end{aligned}$$

and let $\bar{\mathcal{A}}_b = \cup_{n=1}^\infty \bar{\mathcal{A}}_{b,n}$. Define $\tilde{\mathcal{U}}^N = \mathcal{U}^N \cap \bar{\mathcal{A}}_b$.

Theorem 4.2.1. *Suppose Condition 2.3.2 holds with $\mathcal{U}^N$ replaced by $\tilde{\mathcal{U}}^N$. Then the conclusions of Theorem 2.3.2 continue to hold.*

4.2.2 A family of SPDEs driven by Poisson Random Measures.

In this section we introduce the basic SPDE model that will be studied in this work. We begin by giving a precise meaning to a solution for such a SPDE and then recall a result from [52] which gives sufficient conditions on the coefficients ensuring the strong existence and pathwise uniqueness of solutions.

If Φ is a countable Hilbertian nuclear space (recall the definition of CHNS in 3.2.2), and $\{\Phi_n\}_{n\in\mathbb{N}_0}$ are the completions of Φ with respect to $\{||\cdot||_n\}_{n\in\mathbb{N}_0}$, respectively, then $\{\Phi_n\}_{n\in\mathbb{N}_0}$ is a sequence of decreasing Hilbert spaces and $\Phi = \cap_{n=0}^\infty \Phi_n$. Recall, we identify Φ_0' with Φ_0 using Riesz's representation theorem, and denote the space of bounded linear functionals on Φ_n by Φ_{-n}. This space has a natural inner product (and norm) which we denote by $\langle \cdot, \cdot \rangle_{-n}$ (resp. $|| \cdot ||_{-n}$), $n \in \mathbb{N}_0$ such that $\{\Phi_{-n}\}_{n \in \mathbb{N}_0}$ is a sequence of increasing Hilbert spaces and the topological dual of Φ, denoted as Φ' equals $\cup_{n=0}^{\infty} \Phi_{-n}$ (see Theorem 1.3.1 of [52]). Elements of Φ' need not have much regularity. Solutions of the SPDE considered in this chapter will have sample paths in Φ'. In fact under the conditions imposed here the solutions will take values in $D([0,T] : \Phi_{-n})$ for some finite value of n.

We will assume that there is a sequence $\{\phi_j\} \subset \Phi$ such that $\{\phi_j\}$ is a complete orthonormal system (CONS) in Φ_0 and is a complete orthogonal system (COS) in each $\Phi_n, n \in \mathbb{Z}$. Then $\{\phi_j^n\} = \{\phi_j ||\phi_j||_n^{-1}\}$ is a CONS in Φ_n for each $n \in \mathbb{Z}$. Define the map $\theta_p : \Phi_{-p} \to \Phi_p$ by $\theta_p(\phi_j^{-p}) = \phi_j^p$. It is easy to check that for all $p \in \mathbb{N}$, $\theta_p(\Phi) \subseteq \Phi$ (see Remark 6.1.1 of [52]). Also, for each $r > 0$, $\eta \in \Phi_{-r}$ and $\phi \in \Phi_r$, $\eta[\phi]$ is defined by the formula

$$\eta[\phi] = \sum_{j=1}^{\infty} \langle \eta, \phi_j \rangle_{-r} \langle \phi, \phi_j \rangle_r. \tag{4.2.2}$$

We refer the reader to Example 1.3.2 of [52] for a canonical example of such a Countable Hilbertian Nuclear Space (CHNS) defined using a closed densely defined self-adjoint operator on Φ_0. A similar example was considered in Section 3.2.

Following [51], we introduce the following conditions on the coefficients A and G in equation (4.1.1). Let $A : [0,T] \times \Phi' \to \Phi'$, $G : [0,T] \times \Phi' \times \mathbb{X} \to \Phi'$ be maps satisfying the following condition.

Condition 4.2.1. *There exists $p_0 \in \mathbb{N}$ such that, for every $p \geq p_0$, there exists $q \geq p$ and a constant $K = K(p,q)$ such that the following hold.*

1. *(Continuity) For all $t \in [0,T]$ and $u \in \Phi_{-p}$, $A(t,u) \in \Phi_{-q}$ and $G(t,u,\cdot) \in L^2(\mathbb{X}, \nu; \Phi_{-p})$. The maps $u \mapsto A(t,u)$ and $u \mapsto G(t,u,\cdot)$ are continuous.*

2. *(Coercivity) For all* $t \in [0,T]$*, and* $\phi \in \Phi$*,*

$$2A(t,\phi)[\theta_p \phi] \leq K(1 + ||\phi||_{-p}^2).$$

3. *(Growth) For all* $t \in [0,T]$*, and* $u \in \Phi_{-p}$*,*

$$||A(t,u)||_{-q}^2 \leq K(1 + ||u||_{-p}^2)$$

and

$$\int_{\mathbb{X}} ||G(t,u,v)||_{-p}^2 \nu(dv) \leq K(1 + ||u||_{-p}^2).$$

4. *(Monotonicity) For all* $t \in [0,T]$*, and* $u_1, u_2 \in \Phi_{-p}$*,*

$$\begin{aligned} &2\langle A(t,u_1) - A(t,u_2), u_1 - u_2 \rangle_{-q} \\ &\quad + \int_{\mathbb{X}} ||G(t,u_1,v) - G(t,u_2,v)||_{-q}^2 \nu(dv) \leq K||u_1 - u_2||_{-q}^2. \end{aligned}$$

We now give a precise definition of a solution to the SDE (4.1.1).

Definition 4.2.1. Let $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}), \bar{\mathbb{P}}, \{\bar{\mathcal{F}}_t\})$ be the filtered probability space from Section 2.2.2. Fix $p \in \mathbb{N}_0$, suppose that X_0 is a $\bar{\mathcal{F}}_0$-measurable Φ_{-p}-valued random variable such that $\mathbb{E}||X_0||_{-p}^2 < \infty$. A stochastic process $\{X_t^\epsilon\}_{t\in[0,T]}$ defined on $\bar{\mathbb{M}}$ is said to be a Φ_{-p}-valued strong solution to the SDE (4.1.1) with initial value X_0, if

(a) X_t^ϵ is a Φ_{-p}-valued $\bar{\mathcal{F}}_t$-measurable random variable for all $t \in [0,T]$;

(b) $X^\epsilon \in D([0,T] : \Phi_{-p})$ a.s.;

(c) there is a $q \geq p$ such that for all $t \in [0,T]$ and $u \in \Phi_{-p}$, $A(t,u) \in \Phi_{-q}$ and $G(t,u,\cdot) \in L^2(\mathbb{X}, \nu; \Phi_{-q})$, and there exists a sequence $\{\sigma_n\}_{n\geq 1}$ of $\{\bar{\mathcal{F}}_t\}$-stopping times increasing to infinity such that for each $n \geq 1$,

$$\bar{\mathbb{E}} \int_0^{T\wedge\sigma_n} \int_{\mathbb{X}} ||G(s, X_s^\epsilon, v)||_{-q}^2 \nu(dv) ds < \infty$$

and

$$\bar{\mathbb{E}}\int_0^{T\wedge\sigma_n} ||A(s,X_s^\epsilon)||_{-q}^2 ds < \infty;$$

(d) for all $t \in [0,T]$, almost all $\omega \in \bar{\mathbb{M}}$, and all $\phi \in \Phi$

$$X_t^\epsilon[\phi] = X_0[\phi] + \int_0^t A(s,X_s^\epsilon)[\phi]ds + \epsilon\int_0^t\int_{\mathbb{X}} G(s,X_{s-}^\epsilon,v)[\phi]\tilde{N}^{\epsilon^{-1}}(dsdv). \tag{4.2.3}$$

In Definition 4.2.1, $\tilde{N}^{\epsilon^{-1}}$ is the compensated version of $N^{\epsilon^{-1}}$ as defined below (4.1.1), with $N^{\epsilon^{-1}}$ having jump rates that are scaled by $1/\epsilon$ and is constructed from $\bar{N}$, as below (2.2.1).

One can similarly define a Φ_{-p}-valued strong solution on an arbitrary filtered probability space supporting a suitable PRM.

Definition 4.2.2 (pathwise uniqueness)**.** We say that the Φ_{-p}-valued solution for the SDE (4.1.1) has the **pathwise uniqueness** property if the following is true. Suppose that X and X' are two Φ_{-p}-valued solutions defined on the same filtered probability space with respect to the same Poisson random measure and starting from the same initial condition X_0. Then the paths of X and X' coincide for almost all ω.

The following theorem is taken from [52] (see Theorem 6.2.2, Lemma 6.3.1 and Theorem 6.3.1 therein).

Theorem 4.2.2. *Suppose that Condition 4.2.1 holds. Let X_0 be a Φ_{-p}-valued random variable satisfying $\mathbb{E}||X_0||_{-p}^2 < \infty$. Then for sufficiently large $p_1 \geq p$, the canonical injection from Φ_{-p} to Φ_{-p_1} is Hilbert-Schmidt, and for all such p_1 the SDE* (4.1.1) *with initial value X_0 has a pathwise unique Φ_{-p_1}-valued strong solution.*

4.3 Large Deviation Principle.

Throughout this section we will assume that Condition 4.2.1 holds.

Fix $p \geq p_0$ and $X_0 \in \Phi_{-p}$. Let X^ϵ be the Φ_{-p_1}-valued strong solution to the SDE (4.1.1) with initial value X_0. In this section, we establish an LDP for $\{X^\epsilon\}$ under suitable assumptions, by verifying the sufficient condition in Section 4.2.1.

We begin by introducing the map $\mathcal{G}^0$ that will be used to define the rate function and also used for verification of Condition 2.3.2. Recall that $\mathbb{S} = \cup_{N\geq 1} S^N$, where S^N is defined in (2.3.1). As a first step we show that under Conditions 4.3.1 and 4.3.2 below, for every $g \in \mathbb{S}$, the integral equation

$$\tilde{X}_t^g = X_0 + \int_0^t A(s, \tilde{X}_s^g)ds + \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_s^g, v)(g(s,v) - 1)\nu(dv)ds \tag{4.3.1}$$

has a unique continuous solution. Here g plays the role of a control. Keeping in mind that (4.2.3) is driven by the compensated measure and that equations such as (4.3.1) will arise as law of large number limits, g corresponds to a shift in the scaled jump rate away from that of the original model, which corresponds to $g = 1$. Let

$$||G(t,v)||_{0,-p} = \sup_{u\in\Phi_{-p}} \frac{||G(t,u,v)||_{-p}}{1 + ||u||_{-p}}, \quad (t,v) \in [0,T] \times \mathbb{X}.$$

Condition 4.3.1 (Exponential Integrability). *There exists $\delta_1 \in (0,\infty)$ such that for all $E \in \mathcal{B}([0,T] \times \mathbb{X})$ satisfying $\nu_T(E) < \infty$,*

$$\int_E e^{\delta_1 ||G(s,v)||_{0,-p}^2} \nu(dv)ds < \infty.$$

Remark 4.3.1. Under Condition 4.3.1, for every $\delta_2 \in (0,\infty)$ and for all $E \in \mathcal{B}([0,T] \times \mathbb{X})$ satisfying $\nu_T(E) < \infty$

$$\int_E e^{\delta_2 ||G(s,v)||_{0,-p}} \nu(dv)ds < \infty.$$

The proof of Remark 4.3.1 is given in the appendix.

Remark 4.3.2. The following inequalities will be used several times. Proofs are omitted.

1. For $a, b \in (0, \infty)$, $\sigma \in (1, \infty)$

$$ab \leq e^{\sigma a} + \frac{1}{\sigma}(b \log b - b + 1) = e^{\sigma a} + \frac{1}{\sigma} l(b). \tag{4.3.2}$$

2. For each $\beta > 0$ there exists $c_1(\beta) > 0$, such that $c_1(\beta) \to 0$ as $\beta \to \infty$ and

$$|x - 1| \leq c_1(\beta) l(x) \text{ whenever } |x - 1| \geq \beta.$$

3. For each $\beta > 0$ there exists $c_2(\beta) < \infty$, such that

$$|x - 1|^2 \leq c_2(\beta) l(x) \text{ whenever } |x - 1| \leq \beta.$$

In particular, using the inequalities we have the following lemma.

Lemma 4.3.1. *Under Conditions 4.2.1 (c) and 4.3.1, for every $M \in \mathbb{N}$,*

$$\sup_{g \in S^M} \int_{\mathbb{X}_T} ||G(s, v)||_{0,-p}^2 (g(s, v) + 1) \nu(dv) ds < \infty, \tag{4.3.3}$$

$$\sup_{g \in S^M} \int_{\mathbb{X}_T} ||G(s, v)||_{0,-p} |g(s, v) - 1| \nu(dv) ds < \infty. \tag{4.3.4}$$

and

$$\lim_{\delta \to 0} \sup_{g \in S^M} \sup_{|t-s| \leq \delta} \int_{[s,t] \times \mathbb{X}} ||G(r, v)||_{0,-p} |g(r, v) - 1| \nu(dv) dr = 0. \tag{4.3.5}$$

Proof. First notice that under Condition 4.2.1 (c), we have

$$\int_{\mathbb{X}_T} ||G(s, v)||_{0,-p}^2 \nu(dv) ds \leq KT < \infty. \tag{4.3.6}$$

Thus we only need to prove that

$$\sup_{g \in S^M} \int_{\mathbb{X}_T} ||G(s, v)||_{0,-p}^2 g(s, v) \nu(dv) ds < \infty.$$

If $E = \{(s, v) : ||G(s, v)||_{0,-p} \geq 1\}$, then by (4.3.6) we have $\nu_T(E) < \infty$. Also, from the super linear growth of the function l, we can find κ_1, $\kappa_2 \in (0, \infty)$ such that for all $x \geq \kappa_1$, $x \leq \kappa_2 l(x)$. Define $F = \{(s,v) : g(s,v) \geq \kappa_1\}$. Then, from (4.3.2)

$$\begin{aligned}
&\int_{\mathbb{X}_T} ||G(s,v)||_{0,-p}^2 g(s,v)\nu(dv)ds \\
&\qquad = \int_E ||G(s,v)||_{0,-p}^2 g(s,v)\nu(dv)ds + \int_{E^c} ||G(s,v)||_{0,-p}^2 g(s,v)\nu(dv)ds \\
&\qquad \leq \int_E e^{\delta_1 ||G(s,v)||_{0,-p}^2}\nu(dv)ds + \int_E l(\frac{g(s,v)}{\delta_1})\nu(dv)ds \\
&\qquad\quad + \int_{E^c\cap F} \kappa_2 l(g(s,v))\nu(dv)ds + \kappa_1 \int_{E^c\cap F^c} ||G(s,v)||_{0,-p}^2 \nu(dv)ds.
\end{aligned}$$

Combining this estimate with Condition 4.3.1 and the definition of S^M, we have (4.3.3).

We now prove (4.3.4) and (4.3.5). Note that

$$\begin{aligned}
&\int_{[s,t]\times\mathbb{X}} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr \\
&\qquad = \int_{([s,t]\times\mathbb{X})\cap E} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr \\
&\qquad\quad + \int_{([s,t]\times\mathbb{X})\cap E^c} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr.
\end{aligned}$$

Using (4.3.2) twice (once with $b = g$ and once with $b = 1$), for any $M_0 \in (1,\infty)$

$$\int_{([s,t]\times\mathbb{X})\cap E} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr \leq 2\int_{([s,t]\times\mathbb{X})\cap E} e^{M_0||G(r,v)||_{0,-p}}\nu(dv)dr + \frac{M}{M_0}. \tag{4.3.7}$$

Recalling Remark 4.3.2, for any $\theta > 0$ and $g \in S^M$

$$\begin{aligned}
&\int_{([s,t]\times\mathbb{X})\cap E^c} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr \\
&\qquad = \int_{([s,t]\times\mathbb{X})\cap E^c\cap\{|g-1|\leq\theta\}} ||G(r,v)||_{0,-p}|g-1|\nu(dv)dr \\
&\qquad\quad + \int_{([s,t]\times\mathbb{X})\cap E^c\cap\{|g-1|>\theta\}} ||G(r,v)||_{0,-p}|g-1|\nu(dv)dr \\
&\qquad \leq \left(\int_{[s,t]\times\mathbb{X}} ||G(r,v)||_{0,-p}^2\nu(dv)dr\right)^{1/2}\sqrt{c_2(\theta)M} + c_1(\theta)M. \qquad (4.3.8)
\end{aligned}$$

The inequality in (4.3.4) now follows on setting $s = 0$, $t = T$ in (4.3.7) and (4.3.8) and using Condition 4.2.1 (c) and Remark 4.3.1.

Next consider (4.3.5). Fix $\epsilon \in (0,\infty)$. Choose M_0 such that $\frac{M}{M_0} \leq \frac{\epsilon}{4}$. Let $\delta_1 \in (0,\infty)$ be such that

$$2 \sup_{|t-s|\leq\delta_1} \int_{([s,t]\times\mathbb{X})\cap E} e^{M_0||G(r,v)||_{0,-p}} \nu(dv)dr \leq \frac{\epsilon}{4}.$$

Now choose $\theta \in (0,\infty)$ such that $c_1(\theta)M \leq \frac{\epsilon}{4}$. Finally, choose $\delta_2 \in (0,\infty)$ such that

$$\sup_{|t-s|\leq\delta_2} \left(\int_{[s,t]\times\mathbb{X}} ||G(r,v)||_{0,-p}^2 \nu(dv)dr\right)^{1/2} \sqrt{c_2(\theta)N} \leq \frac{\epsilon}{4}.$$

Using the above inequalities in (4.3.7) and (4.3.8), we have for all $\delta \leq \min\{\delta_1,\delta_2\}$,

$$\sup_{g\in S^M} \sup_{|t-s|\leq\delta} \int_{[s,t]\times\mathbb{X}} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr \leq \epsilon$$

The result follows. □

We will need the following stronger condition on fluctuations of G than (d) of Condition 4.2.1. Let

$$||G(t,v)||_{1,-q} = \sup_{u_1,u_2\in\Phi_{-q},u_1\neq u_2} \frac{||G(t,u_1,v)-G(t,u_2,v)||_{-q}}{||u_1-u_2||_{-q}}.$$

Condition 4.3.2. *For q as in Condition 4.2.1, there exists $\delta \in (0,\infty)$ such that for all $E \in \mathcal{B}([0,T]\times\mathbb{X})$ satisfying $\nu_T(E) < \infty$,*

$$\int_E e^{\delta||G(s,v)||_{1,-q}^2} \nu(dv)ds < \infty.$$

Remark 4.3.3. Under Conditions 4.2.1 (d) and 4.3.2, for every $M \in \mathbb{N}$,

$$\sup_{g\in S^M} \int_{\mathbb{X}_T} ||G(s,v)||_{1,-q}^2 (g(s,v)+1)\nu(dv)ds < \infty,$$

and

$$\sup_{g\in S^M} \int_{\mathbb{X}_T} ||G(s,v)||_{1,-q}|g(s,v)-1|\nu(dv)ds < \infty. \tag{4.3.9}$$

The proof of this remark is similar to that of Lemma 4.3.1, and thus omitted. Note that Conditions 4.3.1 and 4.3.2 hold trivially if $||G(s,v)||_{0,-p}$ and $||G(s,v)||_{1,-q}$ are bounded in (s,v).

Recall that $p_1 \geq p$ is chosen such that the canonical injection from Φ_{-p} to Φ_{-p_1} is Hilbert-Schmidt.

Theorem 4.3.1. *Fix $g \in \mathbb{S}$. Suppose Conditions 4.2.1, 4.3.1 and 4.3.2 hold, and that $X_0 \in \Phi_{-p}$. Then there exists a unique $\tilde{X}^g \in C([0,T] : \Phi_{-p_1})$ such that for every $\phi \in \Phi$,*

$$\tilde{X}^g_t[\phi] = X_0[\phi] + \int_0^t A(s, \tilde{X}^g_s)[\phi]ds + \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}^g_s, v)[\phi](g(s,v)-1)\nu(dv)ds. \quad (4.3.10)$$

Furthermore, for $N \in \mathbb{N}$, $\sup_{t\in[0,T]} \sup_{g \in S^N} ||\tilde{X}^g_t||_{-p} < \infty$.

We note that in the above theorem $\tilde{X}^g$ is a non-random element of $C([0,T] : \Phi_{-p_1})$. We can now present the main large deviations result. Recall that for $g \in \mathbb{S}$, $\nu^g_T(dsdv) = g(s,v)\nu(dv)ds$. Define

$$\mathcal{G}^0(\nu^g_T) = \tilde{X}^g \text{ for } g \in \mathbb{S}, \text{ with } \tilde{X}^g \text{ given by (4.3.10).} \quad (4.3.11)$$

Let $I : D([0,T] : \Phi_{-p_1}) \to [0,\infty]$ be defined as in (4.2.1).

Theorem 4.3.2. *Suppose that Conditions 4.2.1, 4.3.1 and 4.3.2 hold. Then I is a rate function on Φ_{-p_1}, and the family $\{X^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle on $D([0,T] : \Phi_{-p_1})$ with rate function I.*

We now proceed with the proofs. In Section 4.3.1 we prove Theorem 4.3.1 and in Section 4.3.2, we present the proof of Theorem 4.3.2.

4.3.1 Proof of Theorem 4.3.1.

The proof of the theorem is based on the following two lemmas. The first lemma is standard and so its proof is relegated to the appendix. The norm $||\cdot||$ in the lemma is the Euclidean norm in $\mathbb{R}^d$.

Lemma 4.3.2. *Let $a, u : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : [0,T] \times \mathbb{R}^d \to \mathbb{R}$ be measurable functions such that, for a.e. $s \in [0,T]$, the maps $y \mapsto a(s,y)$, $y \mapsto b(s,y)$ and $y \mapsto u(s,y)$ are continuous. Further suppose that for some $\kappa \in (0,\infty)$,*

$$||a(s,y)|| + |b(s,y)| \le \kappa(1 + ||y||), \quad \text{for all } s \in [0,T],\ y \in \mathbb{R}^d$$
$$\int_0^T \sup_{y \in \mathbb{R}^d} ||u(s,y)|| ds \le M < \infty.$$

Fix $x_0 \in \mathbb{R}^d$. Then there exists $x \in C([0,T] : \mathbb{R}^d)$ such that x satisfies the integral equation

$$x(t) = x_0 + \int_0^t a(s, x(s)) ds + \int_0^t b(s, x(s)) u(s, x(s)) ds, \tag{4.3.12}$$

and

$$\sup_{t \in [0,T]} ||x(t)|| \le (||x_0|| + \kappa(M+T)) e^{\kappa(M+T)}.$$

Lemma 4.3.3. *Let $\{a^d, g^d\}_{d \in \mathbb{N}}$ be a sequence of maps, $a^d : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $g^d : [0,T] \times \mathbb{R}^d \times \mathbb{X} \to \mathbb{R}^d$, such that the following hold.*

1. *For each $s \in [0,T]$ and $y \in \mathbb{R}^d$, $g^d(s,y,\cdot) \in L^2(\mathbb{X}, \nu; \mathbb{R}^d)$ and for each $s \in [0,T]$, the maps $y \mapsto a^d(s,y)$ and $y \mapsto g^d(s,y,\cdot)$ (from $\mathbb{R}^d$ to $L^2(\mathbb{X}, \nu; \mathbb{R}^d)$) are continuous.*

2. *For some $\kappa \in (0,\infty)$ and all $d \in \mathbb{N}$,*

$$2\langle a^d(s,y), y\rangle \le \kappa(1 + ||y||^2), \quad \forall (s,y) \in [0,T] \times \mathbb{R}^d$$

   and

$$\int_{\mathbb{X}} ||g^d(s,v)||_0^2 \nu(dv) \le \kappa, \quad \forall s \in [0,T],$$

   where $||g^d(s,v)||_0 = \sup_{y \in \mathbb{R}^d} \frac{||g^d(s,y,v)||}{1+||y||}$.

3. *For each $d \in \mathbb{N}$, there exists $\kappa_d \in (0,\infty)$ with*

$$||a^d(s,y)|| \le \kappa_d(1 + ||y||), \quad \forall (s,y) \in [0,T] \times \mathbb{R}^d.$$

4. There is a $\delta_0 \in (0,\infty)$ such that for all $E \in \mathcal{B}([0,T]\times\mathbb{X})$ satisfying $\nu_T(E) < \infty$,

$$\int_E e^{\delta_0 ||g^d(s,v)||_0} \nu(dv)ds < \infty.$$

Then for any $d \in \mathbb{N}$, $\psi \in \mathbb{S}$ and $x_0^d \in \mathbb{R}^d$, the equation

$$x^d(t) = x_0^d + \int_0^t a^d(s, x^d(s))ds + \int_0^t \int_{\mathbb{X}} g^d(s, x^d(s), v)(\psi(s,v) - 1)\nu(dv)ds \quad (4.3.13)$$

has a solution $x^d \in C([0,T] : \mathbb{R}^d)$. Suppose that $\sup_{d\in\mathbb{N}} ||x_0^d||^2 < \infty$. Then for every $M \in (0,\infty)$, there exists a $\tilde{\kappa}_M \in (0,\infty)$ such that

$$\sup_{d\in\mathbb{N}} \sup_{t\in[0,T]} ||x^d(t)||^2 \leq \tilde{\kappa}_M, \text{ whenever } \psi \in S^M.$$

Proof. For each d fixed, equation (4.3.13) is the same as (4.3.12) with the following choices of a, b and u:

$$a(s,y) = a^d(s,y),$$

$$b(s,y) = 1 + ||y||,$$

and

$$u(s,y) = \int_{\mathbb{X}} \frac{g^d(s,y,v)}{1+||y||}(\psi(s,v) - 1)\nu(dv).$$

Thus in order to prove the existence of the solutions to (4.3.13), it suffices to verify conditions in Lemma 4.3.2. The continuity of a, b and first condition in Lemma 4.3.2 are immediate. The proof of the statement

$$y \mapsto u(s,y) \text{ is continuous for a.e. } s \in [0,T] \quad (4.3.14)$$

is given in the appendix. Finally note that

$$\int_0^T \sup_{y\in\mathbb{R}^d} ||u(s,y)||ds \leq \int_0^T \int_{\mathbb{X}} ||g^d(s,v)||_0 |\psi(s,v) - 1|\nu(dv)ds < \infty,$$

where the last inequality follows from conditions (b) and (d) using a similar argument as for (4.3.4). Thus from Lemma 4.3.2, for each $d \in \mathbb{N}$, there exists a $x^d \in C([0,T] :$

$\mathbb{R}^d$) satisfying (4.3.13). Next note that

$$\begin{aligned}
&||x^d(t)||^2 \\
&= ||x_0^d||^2 + 2\int_0^t \left\langle x^d(s), \left(a^d(s, x^d(s)) + \int_{\mathbb{X}} g^d(s, x^d(s), v)(\psi(s,v) - 1)\nu(dv) \right) \right\rangle ds \\
&\leq ||x_0^d||^2 + 2\int_0^t \left\langle x^d(s), a^d(s, x^d(s)) \right\rangle ds \\
&\quad + 2\int_0^t ||x^d(s)|| \int_{\mathbb{X}} ||g^d(s, x^d(s), v)|| \, |\psi(s,v) - 1|\nu(dv)ds \\
&\leq ||x_0^d||^2 + \kappa \int_0^t (1 + ||x^d(s)||^2)ds \\
&\quad + 2\int_0^t ||x^d(s)||(1 + ||x^d(s)||) \int_{\mathbb{X}} ||g^d(s,v)||_0 |\psi(s,v) - 1|\nu(dv)ds.
\end{aligned} \tag{4.3.15}$$

Let

$$f^d(s) = \int_{\mathbb{X}} ||g^d(s,v)||_0 |\psi(s,v) - 1|\nu(dv).$$

Then as before, using (b) and (d), we have that

$$\sup_{\psi \in S^M} \sup_{d \in \mathbb{N}} \int_0^T f^d(s)ds < \infty. \tag{4.3.16}$$

Also, from (4.3.15) and using that $c + c^2 \leq 1 + 2c^2$ for $c \geq 0$,

$$||x^d(t)||^2 \leq \left(||x_0^d||^2 + \kappa T + 2\int_0^T f^d(s)ds \right) + \int_0^t ||x^d(s)||^2 (\kappa + 4f^d(s))ds.$$

Thus, by Gronwall's inequality

$$||x^d(t)||^2 \leq \left(||x_0^d||^2 + \kappa T + 2\int_0^T f^d(s)ds \right) e^{\kappa t + 4\int_0^t f^d(s)ds}.$$

Hence if $\sup_{d \in \mathbb{N}} ||x_0^d||^2 < \infty$, then by (4.3.16)

$$\sup_{\psi \in S^M} \sup_{d \in \mathbb{N}} \sup_{t \in [0,T]} ||x^d(t)||^2 < \infty.$$

The lemma follows. $\square$

We are now ready to prove Theorem 4.3.1.

Proof of Theorem 4.3.1. We first argue the existence of the solutions to (4.3.10). Let $M \in \mathbb{N}$ be such that $g \in S^M$. Recall the CONS $\{\phi_k^p\}$ defined by $\phi_k^p = \phi_k \left\|\phi_k\right\|_p^{-1} \in \Phi_p$ that was introduced in Section 4.2.2. Fix $d \in \mathbb{N}$ and let $\pi : \Phi_{-p} \to \mathbb{R}^d$ be the mapping given by

$$\pi(u)_k = u[\phi_k^p], \quad k = 1, 2, \ldots, d$$

and denote $\pi(X_0)$ by x_0^d. Define $a^d : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $g^d : [0,T] \times \mathbb{R}^d \times \mathbb{X} \to \mathbb{R}^d$ by

$$a^d(s,x)_k = A\left(s, \sum_{j=1}^d x_j \phi_j^{-p}\right)[\phi_k^p]$$

and

$$g^d(s,x,v)_k = G\left(s, \sum_{j=1}^d x_j \phi_j^{-p}, v\right)[\phi_k^p].$$

It is easy to verify that a^d and g^d satisfy the assumptions of Lemma 4.3.3, and therefore there exists $x^d \in C([0,T] : \mathbb{R}^d)$ which satisfies (4.3.13) with ψ replaced by g. Define the Φ_{-p}-valued continuous function X^d, associated with x^d, by

$$X_t^d = \sum_{k=1}^d (x_t^d)_k \phi_k^{-p}.$$

Then with $\tilde{\kappa}_M$ as in Lemma 4.3.3, we have

$$\sup_{d\in\mathbb{N}} \sup_{t\in[0,T]} ||X_t^d||_{-p}^2 \leq \tilde{\kappa}_M. \tag{4.3.17}$$

Recalling the definition of $u[\phi]$ from (4.2.2), let $\gamma^d : \Phi' \to \Phi'$ be a mapping given by

$$\gamma^d u = \sum_{k=1}^d u[\phi_k^p]\phi_k^{-p}.$$

Let, for $d \in \mathbb{N}$, $A^d : [0,T] \times \Phi' \to \Phi'$ and $G^d : [0,T] \times \Phi' \times \mathbb{X} \to \Phi'$ be measurable mappings given by

$$A^d(s,u) = \gamma^d A(s, \gamma^d u) \quad \text{and} \quad G^d(s,u,v) = \gamma^d G(s, \gamma^d u, v).$$

Then X^d solves

$$X_t^d[\phi] = X_0^d[\phi] + \int_0^t A^d(s, X_s^d)[\phi]ds + \int_0^t \int_{\mathbb{X}} G^d(s, X_s^d, v)[\phi](g(s,v)-1)\nu(dv)ds, \ \ \phi \in \Phi.$$

We now argue that for each $\phi \in \Phi$, the family $\{X^d[\phi]\}_{d\in\mathbb{N}}$ is pre-compact in $C([0,T]:\mathbb{R})$. From (4.3.17), we have

$$\sup_d \sup_{t\in[0,T]} |X_t^d[\phi]| \leq \sup_d \sup_{t\in[0,T]} ||X_t^d||_{-p}||\phi||_p \leq \sqrt{\tilde{\kappa}_M}||\phi||_p < \infty. \tag{4.3.18}$$

Now we consider fluctuations of $X^d[\phi]$. For $0 \leq s \leq t \leq T$,

$$\begin{aligned}
&|X_t^d[\phi] - X_s^d[\phi]| \\
&\leq \int_s^t |A^d(r, X_r^d)[\phi]|dr + \int_s^t \int_{\mathbb{X}} |G^d(r, X_r^d, v)[\phi]| \ |g(r,v)-1|\nu(dv)dr \\
&\leq \int_s^t ||A^d(r, X_r^d)||_{-q}||\phi||_q dr + \int_s^t \int_{\mathbb{X}} ||G^d(r, X_r^d, v)||_{-p}||\phi||_p|g(r,v)-1|\nu(dv)dr.
\end{aligned}$$

Also, for $(s,u) \in [0,T]\times\Phi'$

$$\begin{aligned}
||A^d(s,u)||_{-q}^2 &= \left|\left|\sum_{k=1}^d A(s,\gamma^d u)[\phi_k^p]\phi_k^{-p}\right|\right|_{-q}^2 \\
&= \left|\left|\sum_{k=1}^d A(s,\gamma^d u)[\phi_k^q]\phi_k^{-q}\right|\right|_{-q}^2 \\
&= \sum_{k=1}^d \left(A(s,\gamma^d u)[\phi_k^q]\right)^2 \\
&\leq ||A(s,\gamma^d u)||_{-q}^2 \\
&\leq K\left(1 + ||\gamma^d u||_{-p}^2\right) \\
&\leq K\left(1 + ||u||_{-p}^2\right),
\end{aligned}$$

where for the second equality we use the observation

$$u[\phi_j^q]\phi_j^{-q} = u[\phi_j^p]\phi_j^{-p}, \quad \forall u \in \Phi', \ p,q \geq 0,$$

and the last inequality follows on observing that

$$||\gamma^d u||_{-p}^2 \leq ||u||_{-p}^2, \quad \forall p \geq 0.$$

Similarly,

$$
\begin{aligned}
||G^d(s,u,v)||_{-p}^2 &= \left|\left|\sum_{k=1}^{d} G(s,\gamma^d u,v)[\phi_k^p]\phi_k^{-p}\right|\right|_{-p}^2 \\
&= \sum_{k=1}^{d} \left(G(s,\gamma^d u,v)[\phi_k^p]\right)^2 \\
&\leq ||G(s,\gamma^d u,v)||_{-p}^2.
\end{aligned}
$$

Combining the above estimates we have

$$
\begin{aligned}
|X_t^d[\phi] - X_s^d[\phi]| &\leq ||\phi||_q\sqrt{K}\sqrt{1+\tilde{\kappa}_M}(t-s) \\
&\quad + ||\phi||_p(1+\sqrt{\tilde{\kappa}_M})\int_s^t\int_{\mathbb{X}} ||G(r,v)||_{0,-p}|g(r,v)-1|\nu(dv)dr.
\end{aligned}
$$

By Lemma 4.3.1 we now see that

$$
\lim_{\delta\to 0}\sup_{d\in\mathbb{N}}\sup_{|t-s|\leq\delta} |X_t^d[\phi] - X_s^d[\phi]| = 0. \tag{4.3.19}
$$

Combining (4.3.18) and (4.3.19) we now have that the family $\{X^d[\phi]\}$ is pre-compact in $C([0,T]:\mathbb{R})$ for every $\phi\in\Phi$. Combining this with (4.3.17) we have that $\{X^d\}_{d\in\mathbb{N}}$ is pre-compact in $C([0,T]:\Phi_{-p_1})$ (cf. Theorem 2.5.2 in [52]). Let $\tilde{X}$ be any limit point. Then by the dominated convergence theorem and the definitions of A^d and G^d (see Lemma 6.1.6 and Theorem 6.2.2 of [52]), $\tilde{X}$ satisfies the integral equation (4.3.10). Note that the argument also shows that whenever $g\in S^M$, $\sup_{t\in[0,T]}||\tilde{X}_t||_{-p}^2 \leq \tilde{\kappa}_M$.

Next, we argue uniqueness of solutions. Suppose there are two elements $\tilde{X}$ and $\bar{X}$ of $C([0,T]:\Phi_{-p_1})$ such that both satisfy (4.3.10). Then, using Condition 4.2.1 (d),

$$
\begin{aligned}
||\tilde{X}_t - \bar{X}_t||_{-q}^2 &= 2\int_0^t \langle A(s,\tilde{X}_s) - A(s,\bar{X}_s), \tilde{X}_s - \bar{X}_s\rangle_{-q} ds \\
&\quad + 2\int_0^t\int_{\mathbb{X}} \langle G(s,\tilde{X}_s,v) - G(s,\bar{X}_s,v), \tilde{X}_s - \bar{X}_s\rangle_{-q}(g(s,v)-1)\nu(dv)ds \\
&\leq K\int_0^t ||\tilde{X}_s - \bar{X}_s||_{-q}^2 ds \\
&\quad + 2\int_0^t ||\tilde{X}_s - \bar{X}_s||_{-q}^2 \int_{\mathbb{X}} ||G(s,v)||_{1,-q}|g(s,v)-1|\nu(dv)ds.
\end{aligned}
$$

Also, by Remark 4.3.3,

$$\int_0^T \int_{\mathbb{X}} ||G(s,v)||_{1,-q}|g(s,v)-1|\nu(dv)ds < \infty.$$

An application of Gronwall's inequality now shows that $||\tilde{X}_t - \bar{X}_t||_{-q}^2 = 0$ for all $t \in [0,T]$. Uniqueness follows. $\square$

4.3.2 Proof of Theorem 4.3.2.

From Theorem 4.2.2 and by the classical Yamada-Watanabe argument (cf. [49]), for each $\epsilon > 0$, there exists a measurable map $\mathcal{G}^\epsilon : \mathbb{M} \to D([0,T] : \Phi_{-p_1})$ such that, for any PRM $\mathbf{n}^{\epsilon^{-1}}$ on $[0,T] \times \mathbb{X}$ with mean measure $\epsilon^{-1}\lambda_T \otimes \nu$ given on some filtered probability space, $\mathcal{G}^\epsilon(\epsilon \mathbf{n}^{\epsilon^{-1}})$ is the unique Φ_{-p_1} valued strong solution of (4.1.1) (with $\tilde{N}^{\epsilon^{-1}}$ replaced by $\tilde{\mathbf{n}}^{\epsilon^{-1}} = \mathbf{n}^{\epsilon^{-1}} - \epsilon^{-1}\lambda_T \otimes \nu$) with initial value X_0, where p_1 is as in the statement of Theorem 4.2.2. In particular, $X^\epsilon = \mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}})$ is the strong solution of (4.1.1) with initial value X_0 on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}), \bar{\mathbb{P}}, \{\bar{\mathcal{F}}_t\})$. In view of this observation, for proof of Theorem 4.3.2, it suffices to verify Condition 2.3.2.

We begin with the following lemma.

Lemma 4.3.4. *Fix $N \in \mathbb{N}$, and let $g_n, g \in S^N$ be such that $g_n \to g$ as $n \to \infty$. Let $h : [0,T] \times \mathbb{X} \to \mathbb{R}$ be a measurable function such that*

$$\int_{\mathbb{X}_T} |h(s,v)|^2 \nu_T(dvds) < \infty, \tag{4.3.20}$$

and for all $\delta_2 \in (0,\infty)$

$$\int_E e^{\delta_2 |h(s,v)|} \nu_T(dvds) < \infty, \tag{4.3.21}$$

for all $E \in \mathcal{B}([0,T] \times \mathbb{X})$ satisfying $\nu_T(E) < \infty$. Then

$$\int_{\mathbb{X}_T} h(s,v)(g_n(s,v)-1)\nu_T(dvds) \to \int_{\mathbb{X}_T} h(s,v)(g(s,v)-1)\nu_T(dvds) \tag{4.3.22}$$

as $n \to \infty$.

Proof. We first argue that given $\epsilon > 0$, there exists a compact set $K \subset \mathbb{X}$, such that

$$\sup_n \int_{[0,T]\times K^c} |h(s,v)||g_n(s,v) - 1|\nu(dv)ds \le \epsilon. \tag{4.3.23}$$

For each $\beta \in (0,\infty)$ and compact K in $\mathbb{X}$, the left side of (4.3.23) can be bounded by the sum of the following two terms:

$$T_1 = \sup_n \int_{([0,T]\times K^c)\cap\{|g_n-1|>\beta\}} |h(s,v)||g_n(s,v) - 1|\nu(dv)ds,$$

and

$$T_2 = \sup_n \int_{([0,T]\times K^c)\cap\{|g_n-1|\le\beta\}} |h(s,v)||g_n(s,v) - 1|\nu(dv)ds.$$

Consider T_1 first. Then for every $L \in (1,\infty)$

$$\begin{aligned}
T_1 &\le \sup_n \int_{([0,T]\times K^c)\cap\{|g_n-1|>\beta\}\cap\{|h|<1\}} |h(s,v)||g_n(s,v) - 1|\nu(dv)ds \\
&\quad + \sup_n \int_{([0,T]\times K^c)\cap\{|g_n-1|>\beta\}\cap\{|h|\ge1\}} |h(s,v)||g_n(s,v) - 1|\nu(dv)ds \\
&\le \sup_n \int_{([0,T]\times K^c)\cap\{|g_n-1|>\beta\}\cap\{|h|<1\}} |g_n(s,v) - 1|\nu(dv)ds \\
&\quad + 2\int_{([0,T]\times K^c)\cap\{|h|\ge1\}} e^{L|h(s,v)|}\nu(dv)ds + \frac{1}{L}\sup_n \int_{\mathbb{X}_T} l(g_n(s,v))\nu(dv)ds.
\end{aligned}$$

where the inequality uses (4.3.2) twice (with $b = g_n$ and $b = 1$). Using inequality (b) of Remark 4.3.2, the first term on the right side above can be bounded by

$$c_1(\beta)\sup_n \int_{\mathbb{X}_T} l(g_n(s,v))\nu(dv)ds \le c_1(\beta)N.$$

Therefore,

$$T_1 \le c_1(\beta)N + 2\int_{([0,T]\times K^c)\cap\{|h|\ge1\}} e^{L|h(s,v)|}\nu(dv)ds + \frac{1}{L}N.$$

Now choose β sufficiently large so that $c_1(\beta)N \le \epsilon/6$, L be sufficiently large so that $N/L \le \epsilon/6$. Note that from (4.3.20), $\nu_T\{|h| \ge 1\} < \infty$ and so by (4.3.21), $\int_{|h|\ge1} e^{L|h(s,v)|}\nu_T(dvds) < \infty$. Thus we can find a compact set $K_1 \subset \mathbb{X}$ such that

$$2\int_{([0,T]\times K_1^c)\cap\{|h|\ge1\}} e^{L|h(s,v)|}\nu_T(dvds) \le \epsilon/6.$$

With β chosen as above, consider now the term T_2. We have, using the Cauchy-Schwartz Inequality and inequality (c) of Remark 4.3.2, for every compact K,

$$
\begin{aligned}
T_2^2 &\leq \int_{[0,T]\times K^c} |h(s,v)|^2 \nu(dv)ds \times c_2(\beta) \sup_n \int_{\mathbb{X}_T} l(g_n(s,v))\nu(dv)ds \\
&\leq \int_{[0,T]\times K^c} |h(s,v)|^2 \nu(dv)ds \times c_2(\beta) N.
\end{aligned}
$$

By (4.3.20), we can choose a compact set K_2, such that $T_2 \leq \epsilon/2$ with K replaced by K_2. Thus by taking $K = K_1 \cup K_2$, we have on combining the above estimates that $T_1 + T_2 \leq \epsilon$. This proves (4.3.23).

In order to prove (4.3.22), it now suffices to show that, for every compact $K \subset \mathbb{X}$,

$$
\int_{[0,T]\times K} h(s,v)(g_n(s,v)-1)\nu_T(dvds) \to \int_{[0,T]\times K} h(s,v)(g(s,v)-1)\nu_T(dvds). \quad (4.3.24)
$$

Fix a compact $K \subset \mathbb{X}$. From (4.3.20), we have that $\int_{[0,T]\times K} |h(s,v)|\nu_T(dvds) < \infty$. Thus to prove (4.3.24), it suffices to argue

$$
\int_{[0,T]\times K} h(s,v)g_n(s,v)\nu_T(dvds) \to \int_{[0,T]\times K} h(s,v)g(s,v)\nu_T(dvds). \quad (4.3.25)
$$

When h is bounded, (4.3.25) can be established using Lemma 2.8 in [9]. For completeness we include the proof in Appendix. For general h (which may not be bounded), it is enough to show

$$
\sup_n \int_{[0,T]\times K} |h(s,v)| 1_{\{|h|\geq M\}} g_n(s,v)\nu_T(dvds) \to 0, \quad (4.3.26)
$$

as $M \to \infty$. We have

$$
\begin{aligned}
&\sup_n \int_{[0,T]\times K} |h(s,v)| 1_{\{|h|\geq M\}} g_n(s,v)\nu_T(dvds) \\
&\quad \leq \sup_n \int_{([0,T]\times K)\cap\{|h|\geq M\}} e^{L|h(s,v)|}\nu(dv)ds + \frac{1}{L}\sup_n \int_{\mathbb{X}_T} l(g_n(s,v))\nu(dv)ds \\
&\quad \leq \int_{([0,T]\times K)\cap\{|h|\geq M\}} e^{L|h(s,v)|}\nu(dv)ds + \frac{1}{L}N.
\end{aligned}
$$

Given $\epsilon > 0$, we can choose L large enough such that $N/L \leq \epsilon/2$. Also, since

$$
\int_{[0,T]\times K} e^{L|h(s,v)|}\nu_T(dvds) < \infty,
$$

we can choose M_0 large enough such that $\int_{([0,T]\times K)\cap\{|h|\geq M\}} e^{L|h(s,v)|}\nu(dv)ds \leq \epsilon/2$, for all $M \geq M_0$. Thus for all $M \geq M_0$, $\sup_n \int_{[0,T]\times K} |h(s,v)|1_{|h|\geq M} g_n(s,v)\nu_T(dvds) \leq \epsilon$. Since $\epsilon > 0$ is arbitrary, (4.3.26) follows. This proves the result. $\square$

We now proceed to verify the first part of Condition 2.3.2. Recall the map $\mathcal{G}^0$ defined in (4.3.11).

Proposition 4.3.1. *Fix $N \in \mathbb{N}$, and let $g_n, g \in S^N$ be such that $g_n \to g$ as $n \to \infty$. Then*

$$\mathcal{G}^0\left(\nu_T^{g_n}\right) \to \mathcal{G}^0\left(\nu_T^{g}\right).$$

Proof. Let $\tilde{X}^n = \mathcal{G}^0\left(\nu_T^{g_n}\right)$. By Theorem 4.3.1, there exists a constant $\tilde{\kappa} \in (0, \infty)$ such that

$$\sup_n \sup_{t\in[0,T]} ||\tilde{X}_t^n||_{-p} \leq \tilde{\kappa}. \tag{4.3.27}$$

Using similar arguments as in the proof of Theorem 4.3.1 (cf. (4.3.18) and (4.3.19)), we have, for any $\phi \in \Phi$,

$$\sup_n \sup_{t\in[0,T]} |\tilde{X}_t^n[\phi]| < \infty.$$

Also,

$$\begin{aligned} |\tilde{X}_t^n[\phi] - \tilde{X}_s^n[\phi]| &\leq ||\phi||_q \sqrt{K}\sqrt{1+\tilde{\kappa}}(t-s) \\ &\quad + ||\phi||_p(1+\sqrt{\tilde{\kappa}}) \int_s^t \int_{\mathbb{X}} ||G(r,v)||_{0,-p}|g_n(r,v) - 1|\nu(dv)dr. \end{aligned}$$

Using (4.3.5) in Lemma 4.3.1 we now have that

$$\lim_{\delta\to 0} \sup_n \sup_{|t-s|\leq\delta} |\tilde{X}_t^n[\phi] - \tilde{X}_s^n[\phi]| = 0.$$

This proves that the family $\{\tilde{X}_t^n[\phi]\}$ is pre-compact in $C([0,T] : \mathbb{R})$ for every $\phi \in \Phi$.

Combining this with (4.3.27), we have that $\{\tilde{X}^n\}_{n\in\mathbb{N}}$ is pre-compact in $C([0,T] : \Phi_{-p_1})$ (see Theorem 2.5.2 in [52]). Let $\tilde{X}$ be any limit point. An application of the dominated convergence theorem shows that, along the convergent subsequence,

$$\int_0^t A(s, \tilde{X}_s^n)[\phi]ds \to \int_0^t A(s, \tilde{X}_s)[\phi]ds \tag{4.3.28}$$

as $n \to \infty$. Furthermore, using the convergence of $\tilde{X}^n$ to $\tilde{X}$, Condition 4.2.1 (d) and (4.3.9), we have that

$$\int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_s^n, v)[\phi](g_n(s,v) - 1)\nu(dv)ds - \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_s, v)[\phi](g_n(s,v) - 1)\nu(dv)ds \to 0. \tag{4.3.29}$$

Here we have used the inequality

$$\left| G(s, \tilde{X}_s^n, v)[\phi] - G(s, \tilde{X}_s, v)[\phi] \right| \le ||G(s,v)||_{1,-q} \sup_{t \in [0,T]} ||\tilde{X}_s^n - \tilde{X}_s||_{-q}$$

along with inequality (4.3.9) in Remark 4.3.3.

Also, from (4.3.27), we have that for some $\kappa_1 \in (0, \infty)$

$$|G(s, \tilde{X}_s, v)[\phi]| \le \kappa_1 ||G(s,v)||_{0,-p}, \quad \forall (s,v) \in \mathbb{X}_T.$$

Combining this with Condition 4.2.1 (c) and Remark 4.3.1, we now get from Lemma 4.3.4 that, as $n \to \infty$,

$$\int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_s, v)[\phi](g_n(s,v) - 1)\nu(dv)ds \to \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_s, v)[\phi](g(s,v) - 1)\nu(dv)ds. \tag{4.3.30}$$

Combining (4.3.28), (4.3.29) and (4.3.30) we now see that $\tilde{X}$ must satisfy the integral equation (4.3.10) for all $\phi \in \Phi$. In view of unique solvability of (4.3.10) (Theorem 4.3.1), it now follows that $\tilde{X} = \mathcal{G}^0(\nu_T^g)$. The result follows. □

We now proceed to the second part of Condition 2.3.2. As noted in Theorem 4.2.1, it suffices to verify this condition with $\mathcal{U}^M$ replaced with $\tilde{\mathcal{U}}^M$.

Recall from the beginning of this section that $X^\epsilon = \mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}})$ is the strong solution of (4.1.1) with initial value X_0 on $(\bar{\mathbb{M}}, \mathcal{B}(\bar{\mathbb{M}}), \bar{\mathbb{P}}, \{\bar{\mathcal{F}}_t\})$. Let $\varphi_\epsilon \in \tilde{\mathcal{U}}^M$, define $\psi_\epsilon = 1/\varphi_\epsilon$, and recall the definitions of $\bar{N}$ and $\bar{\nu}_T$ from Section 2.2.2. Then it is easy to check (see Theorem III.3.24 of [50], see also Lemma 2.3 of [18]) that

$$\mathcal{E}_t^\epsilon(\psi_\epsilon) =$$
$$\exp\left\{\int_{(0,t]\times\mathbb{X}\times[0,\epsilon^{-1}]} \log(\psi_\epsilon(s,x))\bar{N}(ds\,dx\,dr) + \int_{(0,t]\times\mathbb{X}\times[0,\epsilon^{-1}]} (-\psi_\epsilon(s,x)+1)\,\bar{\nu}_T(ds\,dx\,dr)\right\}$$

is an $\{\bar{\mathcal{F}}_t\}$-martingale. Consequently

$$\mathbb{Q}_T^\epsilon(G) = \int_G \mathcal{E}_t^\epsilon(\psi_\epsilon)d\bar{\mathbb{P}}, \quad \text{for } G \in \mathcal{B}(\bar{\mathbb{M}})$$

defines a probability measure on $\bar{\mathbb{M}}$, and furthermore $\bar{\mathbb{P}}$ and $\mathbb{Q}_T^\epsilon$ are mutually absolutely continuous. Also it can be verified that under $\mathbb{Q}_T^\epsilon$, $\epsilon N^{\epsilon^{-1}\varphi_\epsilon}$ has the same law as that of $\epsilon N^{\epsilon^{-1}}$ under $\bar{\mathbb{P}}$. Thus it follows that $\tilde{X}^\epsilon = \mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}\varphi_\epsilon})$ is the unique solution of the following controlled stochastic differential equation:

$$\tilde{X}_t^\epsilon = X_0 + \int_0^t A(s, \tilde{X}_s^\epsilon)ds + \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_{s-}^\epsilon, v)\left(\epsilon N^{\epsilon^{-1}\varphi_\epsilon}(dsdv) - \nu(dv)ds\right). \quad (4.3.31)$$

Proposition 4.3.2. *Fix $M \in \mathbb{N}$. Let $\varphi_\epsilon, \varphi \in \tilde{\mathcal{U}}^M$ be such that φ_ϵ converges in distribution to φ, under $\bar{\mathbb{P}}$, as $\epsilon \to 0$. Then $\mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}\varphi_\epsilon}) \Rightarrow \mathcal{G}^0(\nu^\varphi)$.*

Proof. If $\tilde{X}^\epsilon = \mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}\varphi_\epsilon})$, then as just noted, $\tilde{X}^\epsilon$ is the unique solution of (4.3.31). We now show that the family $\{\tilde{X}^\epsilon\}_{\epsilon>0}$ of $D([0,T]:\Phi_{-p_1})$ valued random variables is tight.

We begin by showing that for some $\epsilon_0 \in (0,\infty)$

$$\sup_{0<\epsilon<\epsilon_0} \mathbb{E} \sup_{0\le t\le T} ||\tilde{X}_t^\epsilon||_{-p}^2 < \infty. \quad (4.3.32)$$

Recall that θ_p is defined by $\theta_p(\phi_j^{-p}) = \phi_j^p$ for the CONS $\{\phi_j^{-p}, j \in \mathbb{Z}\}$. By Itô's formula,

$$\begin{aligned}
||\tilde{X}^\epsilon_t||^2_{-p} =& ||X_0||^2_{-p} + 2\int_0^t A(s,\tilde{X}^\epsilon_s)[\theta_p \tilde{X}^\epsilon_s]ds \\
&+ 2\int_0^t \int_{\mathbb{X}} \langle G(s,\tilde{X}^\epsilon_s,v), \tilde{X}^\epsilon_s\rangle_{-p}(\varphi_\epsilon - 1)\nu(dv)ds \\
&+ \int_0^t \int_{\mathbb{X}} \Big(||\epsilon G(s,\tilde{X}^\epsilon_{s-},v)||^2_{-p} + 2\langle \epsilon G(s,\tilde{X}^\epsilon_{s-},v), \tilde{X}^\epsilon_{s-}\rangle_{-p}\Big) \\
&\qquad\qquad \times \Big(N^{\epsilon^{-1}\varphi_\epsilon}(dsdv) - \epsilon^{-1}\varphi_\epsilon \nu(dv)ds\Big) \\
&+ \epsilon \int_0^t \int_{\mathbb{X}} ||G(s,\tilde{X}^\epsilon_s,v)||^2_{-p}\varphi_\epsilon \nu(dv)ds.
\end{aligned} \tag{4.3.33}$$

For completeness we include the proof of (4.3.33) in the appendix.

For the second term in (4.3.33), we have by Condition 4.2.1 (b) that

$$2\int_0^t A(s,\tilde{X}^\epsilon_s)[\theta_p \tilde{X}^\epsilon_s]ds \le K\int_0^t (1+||\tilde{X}^\epsilon_s||^2_{-p})ds. \tag{4.3.34}$$

Also, using $a + a^2 \le 1 + 2a^2$ for $a \ge 0$

$$\begin{aligned}
&\left|\int_0^t \int_{\mathbb{X}} \langle G(s,\tilde{X}^\epsilon_s,v), \tilde{X}^\epsilon_s\rangle_{-p}(\varphi_\epsilon - 1)\nu(dv)ds\right| \\
&\quad \le \int_0^t \int_{\mathbb{X}} \frac{||G(s,\tilde{X}^\epsilon_s,v)||_{-p}}{1+||\tilde{X}^\epsilon_s||_{-p}}(1+||\tilde{X}^\epsilon_s||_{-p})||\tilde{X}^\epsilon_s||_{-p}|\varphi_\epsilon - 1|\nu(dv)ds \\
&\quad \le \int_0^t (1+2||\tilde{X}^\epsilon_s||^2_{-p})\left(\int_{\mathbb{X}} ||G(s,v)||_{0,-p}|\varphi_\epsilon - 1|\nu(dv)\right)ds \\
&\quad \le L_1 + 2\int_0^t ||\tilde{X}^\epsilon_s||^2_{-p}\left(\int_{\mathbb{X}} ||G(s,v)||_{0,-p}|\varphi_\epsilon - 1|\nu(dv)\right)ds,
\end{aligned}$$

where $L_1 = \sup_{\varphi\in S^M}\int_{\mathbb{X}_T} ||G(s,v)||_{0,-p}|\varphi - 1|\nu(dv)ds < \infty$, from (4.3.4).

For the last term in (4.3.33), we have

$$\begin{aligned}
&\epsilon \int_0^t \int_{\mathbb{X}} ||G(s,\tilde{X}^\epsilon_s,v)||^2_{-p}\varphi_\epsilon \nu(dv)ds \\
&\qquad = \epsilon \int_0^t \int_{\mathbb{X}} \frac{||G(s,\tilde{X}^\epsilon_s,v)||^2_{-p}}{(1+||\tilde{X}^\epsilon_s||_{-p})^2}(1+||\tilde{X}^\epsilon_s||_{-p})^2\varphi_\epsilon \nu(dv)ds \\
&\qquad \le 2\epsilon \int_0^t (1+||\tilde{X}^\epsilon_s||^2_{-p})\left(\int_{\mathbb{X}} ||G(s,v)||^2_{0,-p}\varphi_\epsilon \nu(dv)\right)ds \\
&\qquad \le 2\epsilon L_2 + 2\epsilon\int_0^t ||\tilde{X}^\epsilon_s||^2_{-p}\left(\int_{\mathbb{X}} ||G(s,v)||^2_{0,-p}\varphi_\epsilon \nu(dv)\right)ds,
\end{aligned}$$

where $L_2 = \sup_{\varphi \in S^M} \int_{\mathbb{X}_T} ||G(s,v)||^2_{0,-p} \varphi \nu(dv) ds < \infty$, from (4.3.3).

We split the martingale term as $M_t = M_t^1 + M_t^2$, where

$$M_t^1 = \int_0^t \int_{\mathbb{X}} ||\epsilon G(s, \tilde{X}^\epsilon_{s-}, v)||^2_{-p} \left(N^{\epsilon^{-1}\varphi_\epsilon}(dsdv) - \epsilon^{-1}\varphi_\epsilon \nu(dv) ds \right),$$

and

$$M_t^2 = \int_0^t \int_{\mathbb{X}} 2\langle \epsilon G(s, \tilde{X}^\epsilon_{s-}, v), \tilde{X}^\epsilon_{s-} \rangle_{-p} \left(N^{\epsilon^{-1}\varphi_\epsilon}(dsdv) - \epsilon^{-1}\varphi_\epsilon \nu(dv) ds \right).$$

We now use the following Gronwall inequality:

If η and $\psi \geq 0$ satisfy $\eta(s) \leq a + \int_0^s \eta(r)\psi(r) dr$ for all $s \in [0,t]$,

then $\eta(t) \leq a e^{\int_0^t \psi(s) ds}$.

Using this inequality, the above estimates, and Lemma 4.3.1, we have that for some constants $L_3, L_4 \in (1, \infty)$,

$$\sup_{0 \leq s \leq t} ||\tilde{X}^\epsilon_s||^2_{-p} \leq L_3 \left(L_4 + \sup_{0 \leq s \leq t} |M_s^1| + \sup_{0 \leq s \leq t} |M_s^2| \right), \tag{4.3.35}$$

for all $\epsilon \in (0,1)$ and $t \in [0,T]$.

For the term M_t^1, we have, for $\epsilon \in (0,1)$

$$\begin{aligned}
&\mathbb{E} \sup_{0 \leq s \leq T} |M_s^1| \\
&\leq \mathbb{E} \left| \int_0^T \int_{\mathbb{X}} ||\epsilon G(s, \tilde{X}^\epsilon_{s-}, v)||^2_{-p} N^{\epsilon^{-1}\varphi_\epsilon}(dsdv) \right| \\
&\qquad + \mathbb{E} \left| \int_0^T \int_{\mathbb{X}} ||\epsilon G(s, \tilde{X}^\epsilon_{s-}, v)||^2_{-p} \epsilon^{-1} \varphi_\epsilon \nu(dv) ds \right| \\
&\leq 2\mathbb{E} \int_0^T \int_{\mathbb{X}} ||\epsilon G(s, \tilde{X}^\epsilon_s, v)||^2_{-p} \epsilon^{-1} \varphi_\epsilon \nu(dv) ds \\
&\leq 4\epsilon \mathbb{E} \int_0^T (1 + ||\tilde{X}^\epsilon_s||^2_{-p}) \left(\int_{\mathbb{X}} ||G(s,v)||^2_{0,-p} \varphi_\epsilon \nu(dv) \right) ds \\
&\leq 4\epsilon \mathbb{E} \int_{\mathbb{X}_T} ||G(s,v)||^2_{0,-p} \varphi_\epsilon \nu(dv) ds + 4\epsilon \mathbb{E} \sup_{0 \leq s \leq T} ||\tilde{X}^\epsilon_s||^2_{-p} \int_{\mathbb{X}_T} ||G(s,v)||^2_{0,-p} \varphi_\epsilon \nu(dv) ds \\
&\leq 4\epsilon L_2 (1 + \mathbb{E} \sup_{0 \leq s \leq T} ||\tilde{X}^\epsilon_s||^2_{-p}).
\end{aligned} \tag{4.3.36}$$

Next consider the term M^2_t. From the Burkholder-Davis-Gundy inequality, we have that

$$\begin{aligned}
&\mathbb{E}\sup_{0\le s\le T}|M^2_s| \\
&\le 4\mathbb{E}[M^2]_T^{1/2} \\
&\le 4\mathbb{E}\left\{\int_0^T\int_{\mathbb{X}}4\epsilon^2\langle G(s,\tilde{X}^\epsilon_{s-},v),\tilde{X}^\epsilon_{s-}\rangle^2_{-p}N^{\epsilon^{-1}\varphi_\epsilon}(dsdv)\right\}^{1/2} \\
&\le 4\mathbb{E}\left\{\int_0^T\int_{\mathbb{X}}4\epsilon^2||G(s,\tilde{X}^\epsilon_{s-},v)||^2_{-p}||\tilde{X}^\epsilon_{s-}||^2_{-p}N^{\epsilon^{-1}\varphi_\epsilon}(dsdv)\right\}^{1/2} \\
&\le 8\mathbb{E}\left\{\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}\int_0^T\int_{\mathbb{X}}\epsilon^2||G(s,\tilde{X}^\epsilon_{s-},v)||^2_{-p}N^{\epsilon^{-1}\varphi_\epsilon}(dsdv)\right\}^{1/2} \\
&\le \frac{1}{8L_3}\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}+128\epsilon^2L_3\mathbb{E}\left(\int_0^T\int_{\mathbb{X}}||G(s,\tilde{X}^\epsilon_{s-},v)||^2_{-p}N^{\epsilon^{-1}\varphi_\epsilon}(dsdv)\right) \\
&= \frac{1}{8L_3}\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}+128\epsilon L_3\mathbb{E}\left(\int_0^T\int_{\mathbb{X}}||G(s,\tilde{X}^\epsilon_{s},v)||^2_{-p}\varphi_\epsilon\nu(dv)ds\right) \\
&\le \frac{1}{8L_3}\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}+256\epsilon L_2L_3(1+\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}).
\end{aligned} \tag{4.3.37}$$

For the fifth inequality, we have used the AM-GM inequality $\sqrt{ab}\le\frac{a}{2}+\frac{b}{2}$ with $a=\frac{1}{32L_3}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}$ and $b=32L_3\epsilon^2\int_0^T\int_{\mathbb{X}}||G(s,\tilde{X}^\epsilon_{s-},v)||^2_{-p}N^{\epsilon^{-1}\varphi_\epsilon}(dsdv)$. Combining (4.3.35), (4.3.36) and (4.3.37) we now have

$$\left(\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}\right)\left(1-4\epsilon L_2L_3-256\epsilon L_2L_3^2-\frac{1}{8}\right)\le L_3L_4+4L_2L_3+256L_2L_3^2.$$

Choose ϵ_0 small enough so that $\max\{4\epsilon_0L_2L_3,256\epsilon_0L_2L_3^2\}\le\frac{1}{8}$. Then for $\epsilon\le\epsilon_0$, we have that

$$\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p}\le\frac{8}{5}(L_3L_4+4L_2L_3+256L_2L_3^2).$$

This proves (4.3.32).

In view of the estimate in (4.3.32), to prove tightness of $\{\tilde{X}^\epsilon\}_{\epsilon\le\epsilon_0}$ in $D([0,T]:\Phi_{-p_1})$, it suffices to show that for all $\phi\in\Phi$, $\{\tilde{X}^\epsilon[\phi]\}_{\epsilon\le\epsilon_0}$ is tight in $D([0,T]:\mathbb{R})$. For the rest of the proof we will only consider $\epsilon\le\epsilon_0$, however we will suppress ϵ_0 from the notation. Fix $\phi \in \Phi$. Let

$$C_t^\epsilon = \int_0^t A(s, \tilde{X}_s^\epsilon)[\phi] ds + \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_s^\epsilon, v)[\phi](\varphi_\epsilon - 1)\nu(dv) ds$$

and

$$M_t^\epsilon = \epsilon \int_0^t \int_{\mathbb{X}} G(s, \tilde{X}_{s-}^\epsilon, v)[\phi] \tilde{N}^{\epsilon^{-1}\varphi_\epsilon}(ds dv).$$

To argue tightness of C^ϵ in $C([0,T]:\mathbb{R})$, it suffices to show (cf. Lemma 6.1.2 of [52]) that for all $\tau > 0$, there exists $\delta = \delta_\tau > 0$ such that

$$\sup_{0\le\epsilon\le\epsilon_0} \mathbb{P}\left(\sup_{0<\beta-\alpha<\delta} |C_\alpha^\epsilon - C_\beta^\epsilon| > \tau \right) < \tau. \tag{4.3.38}$$

Fix $\tau > 0$. Then for arbitrary $\delta > 0$,

$$\begin{aligned}
&\sup_\epsilon \mathbb{P}\left(\sup_{0<\beta-\alpha<\delta} |C_\alpha^\epsilon - C_\beta^\epsilon| > \tau \right) \\
&= \sup_\epsilon \mathbb{P}\left(\sup_{0<\beta-\alpha<\delta} \left| \int_\alpha^\beta A(s, \tilde{X}_s^\epsilon)[\phi] ds + \int_\alpha^\beta \int_{\mathbb{X}} G(s, \tilde{X}_s^\epsilon, v)[\phi](\varphi_\epsilon - 1)\nu(dv) ds \right| > \tau \right) \\
&\le \sup_\epsilon \mathbb{P}\left(\sup_{0<\beta-\alpha<\delta} \left| \int_\alpha^\beta A(s, \tilde{X}_s^\epsilon)[\phi] ds \right| > \frac{\tau}{2} \right) \\
&\quad + \sup_\epsilon \mathbb{P}\left(\sup_{0<\beta-\alpha<\delta} \left| \int_\alpha^\beta \int_{\mathbb{X}} G(s, \tilde{X}_s^\epsilon, v)[\phi](\varphi_\epsilon - 1)\nu(dv) ds \right| > \frac{\tau}{2} \right) \\
&\le \sup_\epsilon \frac{4}{\tau^2} \mathbb{E}\left(\delta^2 \sup_{0\le s\le T} \left| A(s, \tilde{X}_s^\epsilon)[\phi] \right|^2 \right) \\
&\quad + \sup_\epsilon \frac{2}{\tau} \mathbb{E}\left(\sup_{0<\beta-\alpha<\delta} \left| \int_\alpha^\beta \int_{\mathbb{X}} G(s, \tilde{X}_s^\epsilon, v)[\phi](\varphi_\epsilon - 1)\nu(dv) ds \right| \right).
\end{aligned} \tag{4.3.39}$$

From (4.3.32) and Condition 4.2.1 (c), it follows that

$$\sup_\epsilon \mathbb{E}\left(\sup_{0\le s\le T} \left| A(s, \tilde{X}_s^\epsilon)[\phi] \right|^2 \right) < \infty.$$

Thus we can find $\delta_1 > 0$ such that for all $\delta \le \delta_1$, the first term on the last line of (4.3.39) is bounded by $\tau/2$.

Now we consider the second term:

$$\begin{aligned}
&\left|\int_{[\alpha,\beta]\times\mathbb{X}} G(s,\tilde{X}^\epsilon_s,v)[\phi](\varphi_\epsilon-1)\nu(dv)ds\right| \\
&\quad\le ||\phi||_p\left(1+\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||_{-p}\right)\int_{[\alpha,\beta]\times\mathbb{X}}||G(s,v)||_{0,-p}|\varphi_\epsilon-1|\nu(dv)ds \\
&\quad\le ||\phi||_p\left(1+\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||_{-p}\right)\sup_{g\in S^M}\sup_{|t-s|\le\delta}\int_{[s,t]\times\mathbb{X}}||G(s,v)||_{0,-p}|g-1|\nu(dv)ds.
\end{aligned}$$

Then from (4.3.5) in Lemma 4.3.1 and (4.3.32), we can find $\delta_2>0$ such that for all $\delta\le\delta_2$, the second term on the last line of (4.3.39) is bounded by $\tau/2$. By taking $\delta=\min(\delta_1,\delta_2)$, (4.3.38) holds and the tightness of $\{C^\epsilon\}_{\epsilon\le\epsilon_0}$ follows.

Next consider M^ϵ. We have

$$\begin{aligned}
\mathbb{E}\left\langle M^\epsilon\right\rangle_T &= \epsilon\mathbb{E}\int_0^T\int_{\mathbb{X}}(G(s,\tilde{X}^\epsilon_s,v)[\phi])^2\varphi_\epsilon\nu(dv)ds \\
&\le 2\epsilon||\phi||_p(1+\mathbb{E}\sup_{0\le s\le T}||\tilde{X}^\epsilon_s||^2_{-p})\sup_{\varphi\in S^M}\int_{\mathbb{X}_T}||G(s,v)||^2_{0,-p}\varphi\nu(dv)ds.
\end{aligned}\tag{4.3.40}$$

Using Lemma 4.3.1, we have $\mathbb{E}\sup_{0\le s\le T}\left\langle M^\epsilon\right\rangle_s$ goes to 0 as $\epsilon\to0$. Then by Theorem 6.1.1 in [52], for any $\phi\in\Phi$, the sequence of semimartingales $\tilde{X}^\epsilon_t[\phi]=X_0[\phi]+C^\epsilon_t+M^\epsilon_t$ is tight in $D([0,T]:\mathbb{R})$. It then follows from (4.3.32) and Theorem 2.5.2 in [52] that $\{\tilde{X}^\epsilon\}_{\epsilon\le\epsilon_0}$ is tight in $D([0,T]:\Phi_{-p_1})$.

Now choose a subsequence along which $(\tilde{X}^\epsilon,\varphi_\epsilon,M^\epsilon)$ converges in distribution to $(\tilde{X},\tilde{\varphi},0)$. Without loss of generality, we can assume the convergence is almost sure by using the Skorokhod representation theorem. Note that $\tilde{X}^\epsilon$ satisfies the following integral equation

$$\tilde{X}^\epsilon_t[\phi]=X_0[\phi]+\int_0^t A(s,\tilde{X}^\epsilon_s)[\phi]ds+\int_0^t\int_{\mathbb{X}}G(s,\tilde{X}^\epsilon_s,v)[\phi](\varphi_\epsilon-1)\nu(dv)ds+M^\epsilon.$$

Along the lines of Theorem 4.3.1 and Proposition 4.3.1 (see (4.3.28) – (4.3.30)), we see that $\tilde{X}$ must solve

$$\tilde{X}_t[\phi]=X_0[\phi]+\int_0^t A(s,\tilde{X}_s)[\phi]ds+\int_0^t\int_{\mathbb{X}}G(s,\tilde{X}_s,v)[\phi](\tilde{\varphi}-1)\nu(dv)ds.$$

The unique solvability of the above integral equation gives that $\tilde{X} = \mathcal{G}^0\left(\nu^{\tilde{\varphi}}\right)$, thus we have proved part 2 of Condition 2.3.2, i.e., $\mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}\varphi_\epsilon}) \Rightarrow \mathcal{G}^0\left(\nu^{\varphi}\right)$. $\square$

We are now ready to prove the main theorem.

Proof of Theorem 4.3.2. Using Propositions 4.3.1 and 4.3.2, Theorem 4.3.2 is an immediate consequence of Theorem 4.2.1. $\square$

4.4 Appendix.

Proof of Theorem 4.2.1.

Proof. Proof follows by modifying arguments for the lower bound and upper bound in the proof of Theorem 4.2 of [18].

Lower Bound. Following the proof of Theorem 2.8 in [18], it is easy to see that $-\epsilon \log \bar{\mathbb{E}}\left(e^{-\epsilon^{-1}F(Z^\epsilon)}\right)$ is bounded below (actually equal to)

$$\inf_{\varphi \in \tilde{\mathcal{U}}} \bar{\mathbb{E}}\left[L_T(\varphi) + F \circ \mathcal{G}^\epsilon\left(\epsilon N^{\epsilon^{-1}\varphi}\right)\right], \tag{4.4.1}$$

where $\tilde{\mathcal{U}} = \cup_{N\geq 1}\tilde{\mathcal{U}}^N$. The rest of the proof for the lower bound is as in Theorem 4.2 of [18].

Upper Bound. Fix $\delta \in (0,1)$ and $\phi_0 \in \mathbb{U}$ such that

$$I(\phi_0) + F(\phi_0) \leq \inf_{\phi \in \mathbb{U}}(I(\phi) + F(\phi)) + \delta.$$

Choose $g \in \mathbb{S}_{\phi_0}$ such that $L_T(g) \leq I(\phi_0) + \delta$. Note that $g \in \mathbb{S}_{\phi_0}$ implies $\phi_0 = \mathcal{G}^0\left(\nu_T^g\right)$. Define

$$g_n(t,x) = \begin{cases} \left[g(t,x) \vee \dfrac{1}{n}\right] \wedge n & \text{for } x \in K_n, \\ 1 & \text{else.} \end{cases}$$

Then $g_n \in \bar{\mathcal{A}}_{b,n} \subset \bar{\mathcal{A}}_b$. By the monotone convergence theorem, $L_T(g_n) \uparrow L_T(g)$.

Recalling from the proof of the lower bound that $-\epsilon \log \bar{\mathbb{E}}\left(\exp\left(-\epsilon^{-1}F(Z^\epsilon)\right)\right)$ equals the expression in (4.4.1),

$$\begin{aligned} \limsup_{\epsilon\to 0} -\epsilon \log \bar{\mathbb{E}}\left(e^{-\epsilon^{-1}F(Z^\epsilon)}\right) &\leq L_T(g_n) + \limsup_{\epsilon\to 0} \bar{\mathbb{E}}\left[F \circ \mathcal{G}^\epsilon\left(\epsilon N^{\epsilon^{-1}g_n}\right)\right] \\ &\leq L_T(g_n) + F \circ \mathcal{G}^0\left(\nu_T^{g_n}\right), \end{aligned}$$

where the last inequality follows on observing that since $g_n \in \tilde{\mathcal{U}}^N$ for some N, we have by assumption that, for each fixed n, $\mathcal{G}^\epsilon(\epsilon N^{\epsilon^{-1}g_n}) \Rightarrow \mathcal{G}^0\left(\nu_T^{g_n}\right)$, as $\epsilon \to 0$. Sending $n \to \infty$, we have

$$\begin{aligned} \limsup_{\epsilon\to 0} -\epsilon \log \bar{\mathbb{E}}\left(e^{-\epsilon^{-1}F(Z^\epsilon)}\right) &\leq L_T(g) + F \circ \mathcal{G}^0\left(\nu_T^{g}\right) \\ &\leq I(\phi_0) + \delta + F \circ \mathcal{G}^0\left(\nu_T^{g}\right) \\ &= I(\phi_0) + F(\phi_0) + \delta \\ &\leq \inf_{\phi\in\mathbb{U}}(I(\phi) + F(\phi)) + 2\delta. \end{aligned}$$

Since $\delta \in (0,1)$ is arbitrary the desired upper bound follows. This completes the proof of the theorem. □

Proof of Remark 4.3.1.

Proof. Let $E \in \mathcal{B}(\mathbb{X}_T)$ be such that $\nu_T(E) < \infty$. Fix $\delta_2 \in (0,\infty)$, and define $F = \{(s,v) \in \mathbb{X}_T : ||G(s,v)||_{0,-p} > \delta_2/\delta_1\}$. Then

$$\begin{aligned} \int_E e^{\delta_2 ||G(s,v)||_{0,-p}} \nu(dv)ds &= \int_{E\cap F} e^{\delta_2 ||G(s,v)||_{0,-p}} \nu(dv)ds + \int_{E\cap F^c} e^{\delta_2 ||G(s,v)||_{0,-p}} \nu(dv)ds \\ &\leq \int_{E\cap F} e^{\delta_1 ||G(s,v)||^2_{0,-p}} \nu(dv)ds + e^{\delta_2^2/\delta_1} \int_{E\cap F^c} \nu(dv)ds \\ &\leq \int_{E} e^{\delta_1 ||G(s,v)||^2_{0,-p}} \nu(dv)ds + e^{\delta_2^2/\delta_1} \nu_T(E) < \infty. \end{aligned}$$

The remark follows. □

Proof of Lemma 4.3.2.

Proof. The proof proceeds through a standard Picard iteration argument. Define $x^0(t) = x_0$ for all $t \in [0, T]$. Define $x^n(t)$ iteratively as

$$x^n(t) = x_0 + \int_0^t a(s, x^{n-1}(s))ds + \int_0^t b(s, x^{n-1}(s))u(s, x^{n-1}(s))ds, \quad t \in [0, T].$$

Then

$$\begin{aligned}||x^n(t)|| &\le ||x_0|| + \int_0^t ||a(s, x^{n-1}(s))||ds + \int_0^t ||b(s, x^{n-1}(s))u(s, x^{n-1}(s))||ds \\ &\le ||x_0|| + \int_0^t \kappa(1 + ||x^{n-1}(s)||)ds + \int_0^t \kappa(1 + ||x^{n-1}(s)||) \sup_y ||u(s, y)||ds \\ &\le ||x_0|| + \kappa(M + T) + \kappa \int_0^t ||x^{n-1}(s)||(1 + \sup_y ||u(s, y)||)ds.\end{aligned}$$

Let $L = ||x_0|| + \kappa(M + T)$, $\alpha(s) = 1 + \sup_y ||u(s, x)||$, and $\beta(t) = \int_0^t \alpha(s)ds$. Then a recursive argument shows that for all $t \in [0, T]$,

$$||x^n(t)|| \le L + \kappa L\beta(t) + \frac{\kappa^2 L}{2}\beta(t)^2 + \cdots + \frac{\kappa^n L}{n!}\beta(t)^n,$$

and thus

$$\sup_n \sup_{t\in[0,T]} ||x^n(t)|| \le Le^{\kappa\beta(T)} \le Le^{\kappa(M+T)}. \tag{4.4.2}$$

Similarly

$$\begin{aligned}||x^n(t) - x^n(s)|| &\le \int_s^t ||a(s, x^{n-1}(r))||dr + \int_s^t ||b(r, x^{n-1}(r))u(r, x^{n-1}(r))||dr \\ &\le \kappa(1 + Le^{\kappa(M+T)})(t - s) + \kappa(1 + Le^{\kappa(M+T)}) \int_s^t \sup_y ||u(r, y)||dr,\end{aligned}$$

and therefore

$$\lim_{\delta\to 0} \sup_n \sup_{|t-s|\le\delta} ||x^n(t) - x^n(s)|| = 0.$$

Together with (4.4.2) shows that the sequence $\{x^n\}$ is pre-compact in $C([0, T] : \mathbb{R}^d)$. Let x be a limit point of some subsequence of $\{x^n\}$. Then using the continuity properties of the functions a, b and u with respect to x and the dominated convergence theorem, it is easy to check that x satisfies (4.3.12). The lemma follows. □

Proof of (4.3.14).

Proof. Let $y_n \to y$, $y_n, y \in \mathbb{R}^d$. We will like to show that $u(s, y_n) \to u(s, y)$ for a.e. $s \in [0, T]$. Note that, since $\psi \in S^M$, $\int_{[0,T]\times\mathbb{X}} l(\psi(s, v))\nu(dv)ds \leq M$. Thus there exists $\mathbb{T}_1 \subset [0, T]$, with $\lambda_T(\mathbb{T}_1^c) = 0$ and such that

$$\int_{\mathbb{X}} l(\psi(s, v))\nu(dv) < \infty, \quad \forall s \in \mathbb{T}_1.$$

Also, from arguments similar to those in the proof of Lemma 4.3.1,

$$\int_{\mathbb{X}_T} ||g^d(s, v)||_0|\psi(s, v) - 1|\nu(dv)ds < \infty.$$

Consequently, there exists $\mathbb{T}_2 \subset [0, T]$, with $\lambda_T(\mathbb{T}_2^c) = 0$ and such that

$$\int_{\mathbb{X}} ||g^d(s, v)||_0|\psi(s, v) - 1|\nu(dv) < \infty, \quad \forall s \in \mathbb{T}_2. \tag{4.4.3}$$

Let $\mathbb{T} = \mathbb{T}_1 \cap \mathbb{T}_2$ and fix $s \in \mathbb{T}$. Define $F_\beta(s) = \{v \in \mathbb{X} : |\psi(s, v) - 1| \leq \beta\}$ for $\beta \in (0, \infty)$. Then

$$\begin{aligned} u(s, y_n) &= \int_{\mathbb{X}\cap F_\beta} \frac{g^d(s, y_n, v)}{1 + ||y_n||}(\psi(s, v) - 1)\nu(dv) + \int_{\mathbb{X}\cap F_\beta^c} \frac{g^d(s, y_n, v)}{1 + ||y_n||}(\psi(s, v) - 1)\nu(dv) \\ &= u_1(s, y_n) + u_2(s, y_n). \end{aligned}$$

From part (c) of Remark 4.3.2, for all $v \in F_\beta(s)$,

$$|\psi(s, v) - 1|^2 \leq c_2(\beta)l(\psi(s, v)).$$

Thus $[\psi(s, \cdot) - 1]1_{F_\beta(s)}(\cdot) \in L^2(\mathbb{X}, \nu; \mathbb{R})$. From assumption (a) in Lemma 4.3.3 we now see that, for all such s, $u_1(s, y_n) \to u_1(s, y)$, as $n \to \infty$.

For $u_2(s, y_n)$, we have

$$\left| \frac{g^d(s, y_n, v)}{1 + ||y_n||}(\psi(s, v) - 1) \right| \leq ||g^d(s, v)||_0|\psi(s, v) - 1|.$$

From (4.4.3), the term on the right hand side is ν-integrable. Furthermore, $\nu(F_\beta^c) \to 0$ from the super linear growth of l. Thus $u_2(s, y_n)$ converges to 0, uniformly in n, as β goes to ∞. The term $u_2(s, y)$ can be treated in a similar manner. Thus we have shown that, for all $s \in \mathbb{T}$, $u(s, y_n) \to u(s, y)$. Since $\lambda_T(\mathbb{T}^c) = 0$, the result follows. $\square$

Proof of (4.3.25) when h is a bounded and measurable function.

Proof. We can assume without loss of generality that $\int_K g\nu_T(dsdv) \neq 0$ and $\int_K g_n\nu_T(dsdv) \neq 0$, for all $n \geq 1$. Define probability measures $\tilde{\nu}^n$ and $\tilde{\nu}$ as follows:

$$\tilde{\nu}^n(\cdot) = \frac{\nu_T^{g^n}(\cdot \cap K)}{m_n}, \quad \tilde{\nu}(\cdot) = \frac{\nu_T^g(\cdot \cap K)}{m}$$

where $m_n = \int_K g_n\nu_T(dsdv)$ and $m = \int_K g\nu_T(dsdv)$. If $\theta(\cdot) = \frac{\nu_T(\cdot \cap K)}{\nu_T(K)}$, then θ is also a probability measure. We have

$$\begin{aligned} R(\tilde{\nu}^n||\theta) &= \int_K \log\left(\frac{\nu_T(K)}{m_n}g_n\right)\frac{1}{m_n}g_n\nu_T(dsdv) \\ &= \frac{1}{m_n}\int_K (l(g_n) + g_n - 1)\nu_T(dsdv) + \log\frac{\nu_T(K)}{m_n} \\ &\leq \frac{N}{m_n} + 1 - \frac{\nu_T(K)}{m_n} + \log\frac{\nu_T(K)}{m_n}. \end{aligned}$$

Noting that $m_n \to m$, we have that there exists constant α such that $\sup_{n\in\mathbb{N}} R(\tilde{\nu}^n||\theta) \leq \alpha < \infty$. Also note that $\tilde{\nu}^n$ converges weakly to $\tilde{\nu}$. From Lemma 2.8 of [9], we have

$$\frac{1}{m_n}\int_{[0,T]\times K} h(s,v)g_n(s,v)\nu_T(dvds) \to \frac{1}{m}\int_{[0,T]\times K} h(s,v)g(s,v)\nu_T(dvds),$$

which proves (4.3.25). $\square$

Proof of Itô's formula in (4.3.33).

Proof. Here we will give the proof for a simpler case when X_t satisfies the following integral equation, the proof of (4.3.33) being very similar to this case:

$$X_t = X_0 + \int_0^t A(s,X_s)ds + \int_0^t\int_{\mathbb{X}} G(s,X_{s-},v)\tilde{N}(dsdv).$$

For $j \in \mathbb{N}$,

$$X_t[\theta_p\phi_j] = X_0[\theta_p\phi_j] + \int_0^t A(s,X_s)[\theta_p\phi_j]ds + \int_0^t\int_{\mathbb{X}} G(s,X_{s-},v)[\theta_p\phi_j]\tilde{N}(dsdv).$$

Note that

$$X_t[\theta_p\phi_j] = \langle X_t, \phi_j\rangle_{-p} = ||\phi_j||_{-p}\langle X_t, \phi_j^{-p}\rangle_{-p},$$

so

$$\sum_{j=1}^{\infty} ||\phi_j||_p^2 (X_t[\theta_p\phi_j])^2 = \sum_{j=1}^{\infty}\langle X_t, \phi_j^{-p}\rangle_{-p}^2 = ||X_t||_{-p}^2.$$

If $\xi_j(t) = X_t[\theta_p\phi_j]$, then $\xi_j(t)$ satisfies

$$\xi_j(t) = \xi_j(0) + \int_0^t a^j(s)ds + \int_0^t\int_{\mathbb{X}} b^j(s,v)\tilde{N}(dsdv).$$

where $a^j(s) = A(s, X_s)[\theta_p\phi_j]$ and $b^j(s,v) = G(s, X_{s-}, v)[\theta_p\phi_j]$. Applying Itô's formula (cf. Theorem 2.5.1 of [49]) to the real valued semimartingale $\xi_j(t)$, we have

$$\begin{aligned}\xi_j^2(t) = \xi_j^2(0) &+ 2\int_0^t a^j(s)\xi_j(s)ds + 2\int_0^t\int_{\mathbb{X}} b^j(s,v)\xi_j(s-)\tilde{N}(dsdv)\\ &+ \int_0^t\int_{\mathbb{X}}[b^j(s,v)]^2\tilde{N}(dsdv) + \int_0^t\int_{\mathbb{X}}[b^j(s,v)]^2\nu(dv)ds.\end{aligned} \tag{4.4.4}$$

Note that $||X_t||_{-p}^2 = \sum_{j=1}^{\infty}||\phi_j||_p^2\xi_j^2(t)$. So for the second term in (4.4.4), we have

$$\begin{aligned}\sum_{j=1}^{\infty}||\phi_j||_p^2 a^j(s)\xi_j(s) &= \sum_{j=1}^{\infty}||\phi_j||_p^2 A(s,X_s)[\theta_p\phi_j]X_s[\theta_p\phi_j]\\ &= A(s,X_s)\left[\sum_{j=1}^{\infty}||\phi_j||_p^2 X_s[\theta_p\phi_j]\theta_p\phi_j\right]\\ &= A(s,X_s)\left[\sum_{j=1}^{\infty}||\phi_j||_p^2\langle X_s,\phi_j\rangle_{-p}||\phi_j||_{-p}^2\phi_j\right]\\ &= A(s,X_s)\left[\sum_{j=1}^{\infty}\langle X_s,\phi_j^{-p}\rangle_{-p}\phi_j^p\right]\\ &= A(s,X_s)[\theta_p X_s].\end{aligned}$$

Also, we have

$$\begin{aligned}\sum_{j=1}^{\infty}||\phi_j||_p^2 b^j(s,v)\xi_j(s-) &= \sum_{j=1}^{\infty}||\phi_j||_p^2 G(s,X_{s-},v)[\theta_p\phi_j]X_{s-}[\theta_p\phi_j] \\ &= \sum_{j=1}^{\infty}||\phi_j||_p^2 \langle G(s,X_{s-},v),\phi_j\rangle_{-p}\langle X_{s-},\phi_j\rangle_{-p} \\ &= \sum_{j=1}^{\infty}\langle G(s,X_{s-},v),\phi_j^{-p}\rangle_{-p}\langle X_{s-},\phi_j^{-p}\rangle_{-p} \\ &= \langle G(s,X_{s-},v),X_{s-}\rangle_{-p}.\end{aligned}$$

Finally, notice that

$$\begin{aligned}\sum_{j=1}^{\infty}||\phi_j||_p^2[b^j(s,v)]^2 &= \sum_{j=1}^{\infty}||\phi_j||_p^2\left(G(s,X_{s-},v)[\theta_p\phi_j]\right)^2 \\ &= \sum_{j=1}^{\infty}||\phi_j||_p^2\left(\langle G(s,X_{s-},v),\phi_j\rangle_{-p}\right)^2 \\ &= \sum_{j=1}^{\infty}\left(\langle G(s,X_{s-},v),\phi_j^{-p}\rangle_{-p}\right)^2 \\ &= ||G(s,X_{s-},v)||_{-p}^2.\end{aligned}$$

Combining the above equalities with (4.4.4), we have

$$\begin{aligned}||X_t||_{-p}^2 = ||X_0||_{-p}^2 &+ 2\int_0^t A(s,X_s)[\theta_p X_s]ds + 2\int_0^t\int_{\mathbb{X}}\langle G(s,X_{s-},v),X_{s-}\rangle_{-p}\tilde{N}(dsdv) \\ &+ \int_0^t\int_{\mathbb{X}}||G(s,X_{s-},v)||_{-p}^2\tilde{N}(dsdv) + \int_0^t\int_{\mathbb{X}}||G(s,X_{s-},v)||_{-p}^2\nu(dv)ds.\end{aligned}$$

The result follows. $\square$

Chapter 5

Large Deviations for Degenerate Small Noise Diffusions with Fast Markov Modulated Coefficients

5.1 Introduction.

We consider a two component Markov process (X^ϵ, Y^ϵ) with values in $\mathbb{G} = \mathbb{R}^d \times \mathbb{L}$, where $\epsilon > 0$ is a scaling parameter and $\mathbb{L}$ is a finite set, with infinitesimal generator $\mathcal{L}^\epsilon$ given as

$$\begin{aligned}\mathcal{L}^\epsilon\left(\varphi \otimes \psi\right)(x,y) =& \psi(y)\left(b(x,y)\nabla\varphi(x) + \frac{\epsilon}{2}\mathrm{Tr}(aa^T D^2\varphi)(x)\right) \\ &+\varphi(x)\frac{c(x,y)}{\epsilon}\int_{\mathbb{L}}[\psi(\tilde{y}) - \psi(y)]R(x,y,d\tilde{y}),\end{aligned}$$

where φ is a twice continuously differentiable function on $\mathbb{R}^d$, ψ is a function from $\mathbb{L}$ to $\mathbb{R}$, ∇ is the gradient operator and D^2 is the Hessian matrix. Here a, b, c and R are suitable functions, in particular c is strictly positive and R is a transition probability kernel on $(\mathbb{R}^d \times \mathbb{L}) \times \mathcal{B}(\mathbb{L})$. Roughly speaking, the process X^ϵ between consecutive jumps of Y^ϵ is a diffusion with coefficients $b(x,y)$ and $\sqrt{\epsilon}a(x,y)$, namely denoting by τ the first jump instant of Y^ϵ, on $[0,\tau)$,

$$dX^\epsilon(t) = b(X^\epsilon(t), y_0)dt + \sqrt{\epsilon}a(X^\epsilon(t), y_0)dW(t),$$

where W is a d dimensional Brownian motion and $(X^\epsilon(0), Y^\epsilon(0)) = (x_0, y_0) \in \mathbb{R}^d \times \mathbb{L}$. The process Y^ϵ is a pure jump process with jump intensity function $\epsilon^{-1}c(x,y)$ and transition probability kernel $R(x,y,d\bar{y})$, in particular, with τ as above,

$$\mathbf{P}(\tau > t \mid \sigma\{X^\epsilon(s), s \le t\}) = \exp\left\{-\frac{1}{\epsilon}\int_0^t c(X^\epsilon(s), y_0)ds\right\}$$

and

$$\mathbf{P}\left(Y^\epsilon(\tau) \in d\bar{y} \mid X^\epsilon(\tau-) = x, Y^\epsilon(\tau-) = y_0\right) = R(x, y_0, d\bar{y}).$$

A precise stochastic evolution equation for (X^ϵ, Y^ϵ) will be given in Section 5.2. This pair describes a jump-diffusion, where the diffusion component (i.e. X^ϵ) has "small noise" while the jump component (Y^ϵ) has jumps at rate $O(1/\epsilon)$. Classical Averaging principles [42, 73] show that, under conditions, as $\epsilon \to 0$, X^ϵ converges in probability in $C([0,T]:\mathbb{R}^d)$ (the space of continuous functions from $[0,T]$ to $\mathbb{R}^d$ with the uniform topology), to the solution of an 'averaged' equation

$$\frac{d\xi(t)}{dt} = \hat{b}(\xi(t)), \quad \xi(0) = x_0,$$

where $\hat{b}(x) = \int_{\mathbb{L}} b(x,y)\rho_x(dy)$ and for each $x \in \mathbb{R}^d$, ρ_x is the invariant measure of a $\mathbb{L}$ valued Markov process with jump intensity $c(x,\cdot)$ and transition probability kernel $R(x,\cdot,\cdot)$.

In this work we are interested in the study of a large deviation principle(LDP) for $\{X^\epsilon\}_{\epsilon>0}$, as $\epsilon \to 0$, in $C([0,T]:\mathbb{R}^d)$. Such multiscale systems arise in many applications in engineering, operations research and biological and physical sciences (see [87] and references therein). Large deviation results of the form studied here, in addition to providing probability estimates for non typical events, are a starting point in developing efficient importance sampling algorithms for Monte-Carlo estimation of probabilities of rare events (see [37] and references therein). Large deviation results associated with averaging principles for stochastic dynamical systems have been studied by many authors [39, 40, 42, 32, 78, 58, 80, 79, 60, 56]. The models considered in the current work are usually referred to as systems with "full dependence". This refers to the feature that the coefficients of both the slow and the fast process depend on both variables. Earliest results on large deviations for multiscale systems are due to Freidlin[39, 40], see also [42]. They, in particular, treat the case where the functions c and R do not depend on x (i.e. $c(x,y) \equiv c(y), R(x,y,d\tilde{y}) \equiv R(y,d\tilde{y})$)

and $a = 0$. These papers also consider the case where the two components are given through a system of diffusion equations of the form

$$\begin{aligned} dX^\epsilon(t) =& b_1(X^\epsilon(t), Y^\epsilon(t))dt, \\ dY^\epsilon(t) =& \frac{1}{\epsilon} b_2(X^\epsilon(t), Y^\epsilon(t))dt + \frac{1}{\sqrt{\epsilon}} a_2(Y^\epsilon(t))dW_2(t), \end{aligned} \tag{5.1.1}$$

where W_2 is a Brownian motion and b_1, b_2, a_2 are suitable coefficients. In this case although there is a "full dependence" in the sense described above, the fact that the diffusion coefficient a_2 only depends on the fast variable makes the analysis significantly more tractable since by appealing to Girsanov's theorem one can reduce the problem to a setting where the evolution of the fast variable does not depend on the values of the slow variable. Such a reduction is not possible in the setting considered in the current work. The paper [79] is closer to the setting considered in our work in that, in [79], the equation of the fast variable takes the form

$$dY^\epsilon(t) = \frac{1}{\epsilon} b_2(X^\epsilon(t), Y^\epsilon(t))dt + \frac{1}{\sqrt{\epsilon}} a_2(X^\epsilon(t), Y^\epsilon(t))dW_2(t)$$

and thus a Girsanov transformation that gets rid of full dependence is not possible. However there is a key difference in that, in the current work one has to contend with two forms of asymptotic behavior: small noise effects in the dynamics of X^ϵ; and stochastic averaging effects from the fast variable, whereas in [79] only the latter needs to be understood. Furthermore, we consider here a model where the fast variable is a jump process while [79] considers the setting of diffusions. Finally our proofs are very different from those in [79] which rely on a delicate two level time discretization, whereas the proofs in the current work largely bypass any discretization. Our proofs are based on recent variational representations for functionals of Brownian motions and Poisson random measures, obtained in [18].

Large deviation problems for systems with averaging are closely related to those associated with homogenization problems [41, 3, 36]. In these problems one usually formulates a single equation, with two scaling parameters ϵ and δ, of the form

$$dX^\epsilon(t) = \left[\frac{\epsilon}{\delta}c_1\left(X^\epsilon(t), \frac{X^\epsilon(t)}{\delta}\right) + c_2\left(X^\epsilon(t), \frac{X^\epsilon(t)}{\delta}\right)\right] + \sqrt{\epsilon}a_1\left(X^\epsilon(t), \frac{X^\epsilon(t)}{\delta}\right) dW_1(t),$$

where W_1 is a Brownian motion and a_1, c_1, c_2 are suitable functions. In the special case when $\delta = \epsilon$, this can be rewritten, on defining $Y^\epsilon(t) = \frac{X^\epsilon(t)}{\epsilon}$ and $b_1(x, y) = c_1(x, y) + c_2(x, y)$, as an averaging system of the form

$$\begin{aligned}
dX^\epsilon(t) =& b_1(X^\epsilon(t), Y^\epsilon(t))dt + \sqrt{\epsilon}a_1(X^\epsilon(t), Y^\epsilon(t))dW_1(t),\\
dY^\epsilon(t) =& \frac{1}{\epsilon}b_1(X^\epsilon(t), Y^\epsilon(t))dt + \frac{1}{\sqrt{\epsilon}}a_1(X^\epsilon(t), Y^\epsilon(t))dW_1(t).
\end{aligned}$$

This model is once again a system with 'full dependence' where the diffusion coefficient of the fast variable depends also on the slow variable. However there is one key difference from the models studied in [41, 3, 36] from the systems considered here (apart from the fact that the fast component in our setting is a jump process), namely in all these works the diffusion coefficient a_1 is taken to be uniformly nondegenerate. The main challenge with dealing with degenerate a_1 is obtaining suitable regularity properties of the local rate function (see (5.5.4)) that is used in the proof of the lower bound of the LDP. Roughly speaking our approach is as follows. We regularize the local rate function by adding a small viscosity term in the evolution of the slow system in the form of σB, where B is a standard d dimensional Brownian motion independent of the original driving noise W, and σ is a small parameter. We then first prove a large deviations lower bound for the regularized system and then send $\sigma \to 0$ to recover the lower bound for the original system. Allowing for degenerate diffusion coefficients is important for applications, specially when one considers infinite dimensional noise terms. Although not considered here, we believe that techniques developed here will be useful for obtaining large deviation results for infinite dimensional systems with averaging (see e.g. [22]) as well.

The chapter is organized as follows. In Section 5.2 we introduce the key assumptions and then state our main result (Theorem 5.2.3). Section 5.3 shows that the function $\mathbb{I}$ defined in (5.2.11) is a rate function on $C([0,T]:\mathbb{R}^d)$. In Section 5.4 we prove the large deviation upper bound and in Section 5.5 we prove the lower bound. Theorem 5.2.3 follows on combining the results of Sections 5.3, 5.4 and 5.5. Appendix collects proofs of some auxiliary results.

5.2 Mathematical Preliminaries and Main Result.

We assume without loss of generality that $\mathbb{L}$ is a finite additive group whose zero element is denoted by 0. The following is our first assumption on the coefficients.

Assumption 5.2.1. *(1) c is a bounded measurable map from $\mathbb{G}$ to $[0,\infty)$.*

(2) For each $y, y' \in \mathbb{L}$, $a(\cdot,y), b(\cdot,y)$, $c(\cdot,y)$ and $R(\cdot,y,y')$ are Lipschitz functions from $\mathbb{R}^d$ to $\mathbb{R}^{d\times d}$, $\mathbb{R}^d$, $\mathbb{R}_+$ and $\mathbb{R}_+$, respectively.

Remark 5.2.1. Assumption 5.2.1 (2) in particular says that, for some $\kappa_1 \in (0,\infty)$

$$|b(x,y)| + |a(x,y)| \le \kappa_1(1+|x|), \text{ for all } (x,y) \in \mathbb{G}.$$

Under Assumption 5.2.1 (1), one can construct (see [73] page 104) a finite measure θ on $([0,1],\mathcal{B}[0,1])$ and a measurable function k from $\mathbb{R}^d \times \mathbb{L} \times [0,1]$ to $\mathbb{L}$ such that

$$\begin{aligned} \theta\{r : k(x,y,r) \neq 0\} &= c(x,y), \\ \theta\{r : y + k(x,y,r) \in B\} &= R(x,y,B)c(x,y),\ (x,y) \in \mathbb{G},\ B \subset \mathbb{L}\setminus\{y\}, \end{aligned}$$

where $R(x,y,B) = \sum_{y'\in B} R(x,y,y')$. Denote

$$E_{x,y} = \{r : k(x,y,r) \neq 0\}, \text{ and } E^{y'}_{x,y} = \{r : k(x,y,r)+y = y'\},\ (x,y,y') \in \mathbb{R}^d\times\mathbb{L}\times\mathbb{L}.$$

Then using the boundedness of c and Lipschitz property of c and R, one can assume without loss of generality that θ and k are constructed in a manner that for some

$\kappa_2 \in (0,\infty)$

$$\sup_{(y,y')\in\times\mathbb{L}\times\mathbb{L}} [\theta(E^{y'}_{x,y} \triangle E^{y'}_{x',y}) + \theta(E_{x,y} \triangle E_{x',y})] \le \kappa_2 |x - x'|, \tag{5.2.1}$$

where Δ denotes the symmetric difference operator. Furthermore one can assume without loss of generality that

$$\theta(\partial E^{y'}_{x,y}) + \theta(\partial E_{x,y}) = 0 \quad \forall (x,y,y') \in \mathbb{R}^d \times \mathbb{L} \times \mathbb{L}. \tag{5.2.2}$$

Let $\bar{N}$ be a Poisson random measure (PRM) on $[0,1]\times\mathbb{R}_+\times\mathbb{R}_+$ with intensity measure $\theta\otimes\lambda_\infty\otimes\lambda_\infty$, where λ_∞ denotes the Lebesgue measure on $\mathbb{R}_+$. Then

$$N^{1/\epsilon}(dr\times dt) = \bar{N}(dr \times dt \times [0, \frac{1}{\epsilon}])$$

is a PRM on $[0,1]\times\mathbb{R}_+$ with intensity measure $\frac{1}{\epsilon}\theta\otimes\lambda_\infty$. In terms of this PRM, the evolution of (X^ϵ, Y^ϵ) can be described through the unique pathwise solution of the following system of equations.

State Dynamics.

$$\begin{aligned} dX^\epsilon(t) &= b(X^\epsilon(t), Y^\epsilon(t))dt + \sqrt{\epsilon}a(X^\epsilon(t), Y^\epsilon(t))dW(t),\ X^\epsilon(0) = x_0 && (5.2.3)\\ dY^\epsilon(t) &= \int_{r\in[0,1]} k(X^\epsilon(t), Y^\epsilon(t), r)N^{1/\epsilon}(dr\times dt),\ Y^\epsilon(0) = y_0 && (5.2.4) \end{aligned}$$

where W is a d dimensional Brownian martingale on some complete filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \{\mathcal{F}_t\})$ which also supports the Poisson random measure $\bar{N}$ such that

$$\bar{N}(A\times[0,t]\times B) - t\theta(A)\lambda_\infty(B)$$

is a $\{\mathcal{F}_t\}$ martingale for all $A \in \mathcal{B}[0,1]$ and $B \in \mathcal{B}(\mathbb{R}_+)$ with $\lambda_\infty(B) < \infty$.

For each fixed $x \in \mathbb{R}^d$, the operator Π_x acting on $\mathbb{M}(\mathbb{L})$, defined as

$$\Pi_x\phi(y) = -c(x,y)\phi(y) + c(x,y)\int_{\mathbb{L}} \phi(z)R(x,y,dz),\ y\in\mathbb{L}, \phi\in\mathbb{M}(\mathbb{L}),$$

describes the generator of a $\mathbb{L}$ valued Markov process. Let

$$\hat{R}_n(x,y,z) = \sum_{y' \in \mathbb{L}} \hat{R}(x,y',z)\hat{R}_{n-1}(x,y,y'),\ n > 1;\ \hat{R}_1(x,y,z) = R(x,y,z) \qquad (5.2.5)$$

be the n-step transition probability kernel of the embedded chain. Let $\ell = \mathrm{card}(\mathbb{L})$ and define

$$\alpha_x = \min_{y,z \in \mathbb{L}} \sum_{n=1}^{\ell} \hat{R}_n(x,y,z), \quad \alpha = \inf_x \alpha_x;$$

$$\underline{\lambda}_x = \min_{y \in \mathbb{L}} c(x,y), \quad \underline{\lambda} = \inf_x \underline{\lambda}_x;$$

$$\bar{\lambda}_x = \max_{y \in \mathbb{L}} c(x,y), \quad \bar{\lambda} = \sup_x \bar{\lambda}_x.$$

From Assumption 5.2.1 (1) we see that $\bar{\lambda} < \infty$. Let

$$\mathbb{T} = \{(y,y') \in \mathbb{L} \times \mathbb{L} : R(x,y,y') > 0 \text{ for some } x \in \mathbb{R}^d\}$$

and let

$$\inf_{x \in \mathbb{R}^d} \min_{(y,y') \in \mathbb{T}} R(x,y,y') = \kappa_3.$$

We will make the following ergodicity assumption.

Assumption 5.2.2. *$\alpha > 0$, $\underline{\lambda} > 0$ and $\kappa_3 > 0$.*

Assumptions 5.2.1 and 5.2.2 will be taken to hold throughout this work and will not be mentioned in the statement of various results.

The following is an immediate consequence of our assumptions. For the proof of the second statement in the theorem, see proof of Lemma 5.2.1 in the Appendix.

Theorem 5.2.1. *For each $x \in \mathbb{R}^d$, there is a unique invariant probability measure, ρ_x for the $\mathbb{L}$ valued Markov process with generator Π_x. Furthermore, $\inf_{x \in \mathbb{R}^d} \min_{y \in \mathbb{L}} \rho_x(y) \equiv \underline{\rho} > 0$.*

Define

$$\hat{b}(x) = \int b(x,y)\rho_x(dy).$$

The following lemma ensures that equation (5.2.6) below has a unique solution. Proof of this result is given in the Appendix.

Lemma 5.2.1. *The function $\hat{b}$ is a Lipschitz map on compact subsets of $\mathbb{R}^d$.*

The proof of the following theorem follows as that of Theorem 8 in Chapter 2 of [73].

Theorem 5.2.2. *Fix $(x_0, y_0) \in \mathbb{G}$. Let (X^ϵ, Y^ϵ) solve the system of equations (5.2.3)-(5.2.4). Then as $\epsilon \to 0$, X^ϵ converges uniformly on compacts in probability to the unique solution of*

$$\frac{d\xi(t)}{dt} = \hat{b}(\xi(t)), \quad \xi(0) = x_0 \tag{5.2.6}$$

Let $\mathbb{U} = C([0,T] : \mathbb{R}^d)$. Then the solution X^ϵ of system (5.2.3)-(5.2.4) can be regarded as a $\mathbb{U}$-valued random variable. The main result of this work establishes a large deviation principle (LDP) for X^ϵ, as $\epsilon \to 0$, in the space $\mathbb{U}$. In rest of this section we formulate the rate function for $\{X^\epsilon\}$, and state our main result.

Rate function.

Denote by $\mathcal{M}_F$ the space of all finite measures on $[0,1]$, endowed with the usual topology of weak convergence. For $\eta \in \mathcal{M}_F$ and $x \in \mathbb{R}^d$, consider a $\mathbb{L}$ valued Markov process with infinitesimal generator Π_x^η, defined as

$$\Pi_x^\eta \phi(y) = -\hat{c}^\eta(x,y)\phi(y) + \hat{c}^\eta(x,y)\int_{\mathbb{L}} \phi(z)\hat{R}^\eta(x,y,dz), \; y \in \mathbb{L}, \phi \in \mathbb{M}(\mathbb{L}), \tag{5.2.7}$$

where

$$\begin{aligned} \hat{c}^\eta(x,y) &= \int_{[0,1]} 1_{\{r:k(x,y,r)\neq 0\}}\eta(dr), \\ \hat{c}^\eta(x,y)\hat{R}^\eta(x,y,B) &= \int_{[0,1]} 1_{\{r:y+k(x,y,r)\in B\}}\eta(dr), \; (x,y) \in \mathbb{G}, \; B \subset \mathbb{L} \setminus \{y\}. \end{aligned}$$

We set $\hat{R}^{\eta}(x,y,y)=0$ for all $(x,y)\in\mathbb{G}$. Also, when $\hat{c}^{q}(x,y)=0$, by convention we take $\hat{R}^{\eta}(x,y,y')=R(x,y,y')$. Define $\hat{R}^{\eta}(x,y,A)=\sum_{y\in A}\hat{R}^{\eta}(x,y,y')$ for all $A\subset\mathbb{L}$. Note that with the above notation

$$\Pi_x^{\theta}=\Pi_x,\ \hat{c}^{\theta}(x,y)=c(x,y),\ \hat{R}^{\theta}(x,y,y')=R(x,y,y'),\ (x,y,y')\in\mathbb{R}^d\times\mathbb{L}\times\mathbb{L}.$$

Define $l:[0,\infty)\to[0,\infty)$ as $l(x)=x\log x-x+1$ and let $\hat{l}:\mathcal{M}_F\to[0,\infty]$ be defined as

$$\hat{l}(\eta)=\int_{[0,1]}l(\frac{d\eta}{d\theta})(r)\theta(dr),\ \text{if } \eta<<\theta \text{ and } l(\frac{d\eta}{d\theta}) \text{ is } \theta\text{ -integrable }.$$

Otherwise we set $\hat{l}(\eta)=\infty$.

Denote by $\mathcal{P}_1$ the space of finite measures Q on

$$[0,T]\times\mathbb{L}\times\mathcal{M}_F\times\mathbb{R}^d\equiv\mathbb{H}_T$$

such that

$$Q([a,b]\times\mathbb{L}\times\mathcal{M}_F\times\mathbb{R}^d)=b-a,\quad \text{for all } 0\le a\le b\le T.$$

In other words, denoting the marginal distribution on the i^{th} coordinate of $\mathbb{H}_T$ by $Q_{(i)}$, Q is in $\mathcal{P}_1$ if and only if $Q_{(1)}=\lambda$, where λ is the Lebesgue measure on $[0,T]$. For notational simplicity, we will denote a typical $(s,y,\eta,z)\in\mathbb{H}_T$ as $\mathbf{v}$. For $\xi\in\mathbb{U}$, let $\mathcal{A}_{\xi}$ be the family of all $Q\in\mathcal{P}_1$ such that

$$\int_{\mathbb{H}_T}|z|^2Q(d\mathbf{v})<\infty; \tag{5.2.8}$$

$$\xi(t)=x+\int_{\mathbb{H}_t}b(\xi(s),y)Q(d\mathbf{v})+\int_{\mathbb{H}_t}a(\xi(s),y)zQ(d\mathbf{v}); \tag{5.2.9}$$

and

$$\int_{\mathbb{H}_t}\Pi^{\eta}_{\xi(s)}\phi(y)Q(d\mathbf{v})=0\quad\forall\phi\in\mathbb{M}(\mathbb{L}),\ \forall t\in[0,T], \tag{5.2.10}$$

where $\mathbb{H}_t = [0,t]\times\mathbb{L}\times\mathcal{M}_F[0,1]\times\mathbb{R}^d$. Then the rate function for the family $\{X^\epsilon, \epsilon > 0\}$ is defined to be

$$\mathbb{I}(\xi) = \inf_{Q\in\mathcal{A}_\xi}\left\{\int_{\mathbb{H}_T}\left[\frac{1}{2}|z|^2 + \hat{l}(\eta)\right]Q(d\mathbf{v})\right\}. \tag{5.2.11}$$

The following is the main result of this work. Recall that a function $\mathcal{I} : \mathbb{U} \to [0,\infty]$ is called a rate function on $\mathbb{U}$ if it has compact sub-level sets, namely for every $\alpha \in (0,\infty)$, the set $\{\xi \in \mathbb{U} : \mathcal{I}(\xi) \leq \alpha\}$ is a compact subset of $\mathbb{U}$.

Theorem 5.2.3. *The map $\mathbb{I}$ is a rate function on $\mathbb{U}$ and $\{X^\epsilon\}_{\epsilon>0}$ satisfies a large deviation principle, as $\epsilon \to 0$, on $\mathbb{U}$ with rate function $\mathbb{I}$.*

Rest of the paper is organized as follows. In Section 5.3 we show that $\mathbb{I}$ is a rate function on $\mathbb{U}$. Given this result, to complete the proof of Theorem 5.2.3, it suffices to show that for all $F \in C_b(\mathbb{U})$

$$\lim_{\epsilon\to 0} -\epsilon \log \mathbb{E}\left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right] = \inf_{\xi\in\mathbb{U}}\{F(\xi) + \mathbb{I}(\xi)\}. \tag{5.2.12}$$

In Section 5.4 we show that the left side of (5.2.12) is bounded below by the right side (the upper bound) and in Section 5.5 we prove the reverse inequality. Theorem 5.2.3 follows on combining the results of Sections 5.3, 5.4 and 5.5.

5.3 Compact Level Sets.

In this section we will prove the following result.

Proposition 5.3.1. *For every $M \in (0,\infty)$, the set $\mathbb{U}_M = \{\xi \in \mathbb{U} | \mathbb{I}(\xi) \leq M\}$ is compact and consequently $\mathbb{I}$ is a rate function on $\mathbb{U}$.*

Proof. Let $\{\xi_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{U}_M$. It suffices to show that $\{\xi_n\}$ is pre-compact and every limit point belongs to $\mathbb{U}_M$. Since $\mathbb{I}(\xi_n) \leq M$, we have that for each $n \geq 1$, there exists some $Q_n \in \mathcal{A}_{\xi_n}$, such that

$$\int_{\mathbb{H}_T} \left[\frac{1}{2}|z|^2 + \hat{l}(\eta)\right] Q_n(d\mathbf{v}) \leq M + \frac{1}{n}. \tag{5.3.1}$$

We will now argue that:

(i) $\{Q_n, \xi_n\}_{n\geq 1}$ is pre-compact in $\mathcal{P}_1(\mathbb{H}_T) \times \mathbb{U}$;

(ii) Any limit point $\{Q, \xi\}$ satisfies the following properties.

  (a) $\int_{\mathbb{H}} \left[\frac{1}{2}|z|^2 + \hat{l}(\eta)\right] Q(d\mathbf{v}) \leq M$.

  (b) (5.2.9) holds.

  (c) (5.2.10) holds.

This will show that $Q \in \mathcal{A}_\xi$ and $\mathbb{I}(\xi) \leq M$, completing the proof.

We now give the proofs of (i) and (ii).

Proof of (i). Since $\mathbb{L}$ is a compact set and $(Q_n)_{(1)} = \lambda$ for all n, and $\int_{\mathbb{H}_T} |z|^2 Q_n(d\mathbf{v}) \leq 2(M+1)$, in order to prove the pre-compactness of $\{Q_n\}$, it suffices to show that for every $\delta > 0$, there exists a $c(\delta) \in (0, \infty)$ such that

$$\sup_{n\geq 1} Q_n\{(s, y, \eta, z) \in \mathbb{H}_T | \eta[0,1] > c(\delta)\} \leq \delta. \tag{5.3.2}$$

From superlinearity of l, we see that for some $c_0 \in (0, \infty)$,

$$\eta[0,1] \leq c_0(1 + \hat{l}(\eta)), \quad \forall \eta \in \mathcal{M}_F. \tag{5.3.3}$$

For fixed $\delta > 0$, choosing $c(\delta) \geq \frac{c_0(M+T+1)}{\delta}$, we obtain from Markov's inequality that

$$Q_n\{(s, y, \eta, z) \in \mathbb{H}_T | \eta[0,1] > c(\delta)\} \leq \frac{c_0}{c(\delta)} \left(T + \int_{\mathbb{H}_T} \hat{l}(\eta) Q_n(d\mathbf{v})\right) \leq \delta.$$

This proves (5.3.2), completing the proof of pre-compactness of $\{Q_n\}$.

We next argue the pre-compactness of $\{\xi_n\}$. First, we show that

$$\sup_{n\geq 1}\sup_{0\leq t\leq T}|\xi_n(t)|^2<\infty.$$

Since $Q_n \in \mathcal{A}_{\xi_n}$, we have

$$\xi_n(t)=x+\int_{\mathbb{H}_t} b(\xi_n(s),y)Q_n(d\mathbf{v})+\int_{\mathbb{H}_t} a(\xi_n(s),y)zQ_n(d\mathbf{v}).$$

Using the linear growth condition on a, b, we have

$$\begin{aligned}|\xi_n(t)| \leq & |x|+\int_{\mathbb{H}_t}|b(\xi_n(s),y)|Q_n(d\mathbf{v})+\int_{\mathbb{H}_t}|a(\xi_n(s),y)||z|Q_n(d\mathbf{v})\\ \leq & |x|+\int_{\mathbb{H}_t}\kappa_1(|\xi_n(s)|+1)Q_n(d\mathbf{v})+\sqrt{\int_{\mathbb{H}_t}|a(\xi_n(s),y)|^2Q_n(d\mathbf{v})\int_{\mathbb{H}_t}|z|^2Q_n(d\mathbf{v})}.\end{aligned}$$

Thus

$$\begin{aligned}|\xi_n(t)|^2 \leq & 3|x|^2+6\kappa_1^2\int_{[0,t]}(|\xi_n(s)|^2+1)ds+6(M+1)\int_{\mathbb{H}_t}|a(\xi_n(s),y)|^2Q_n(d\mathbf{v})\\ \leq & 3|x|^2+6\kappa_1^2\int_{[0,t]}(|\xi_n(s)|^2+1)ds+12(M+1)\kappa_1^2\int_{[0,t]}(|\xi_n(s)|^2+1)ds.\end{aligned}$$

Let $A=6\kappa_1^2+12(M+1)\kappa_1^2$ and $B=3|x|^2+6\kappa_1^2T+12(M+1)\kappa_1^2T$. Then we have

$$|\xi_n(t)|^2\leq A\int_{[0,t]}|\xi_n(s)|^2ds+B.$$

By Gronwall's inequality, we have that

$$\sup_{n\geq 1}\sup_{0\leq t\leq T}|\xi_n(t)|^2\leq B\exp(AT)=M_1<\infty.$$

Next, consider fluctuations of ξ_n. We have, for $0\leq t_0\leq t_1\leq T$,

$$\begin{aligned}|\xi_n(t_1)-\xi_n(t_0)| \leq & \int_{[t_0,t_1]\times\mathbb{L}\times\mathcal{M}_F\times\mathbb{R}^d}|b(\xi_n(s),y)|Q_n(d\mathbf{v})\\ & +\int_{[t_0,t_1]\times\mathbb{L}\times\mathcal{M}_F\times\mathbb{R}^d}|a(\xi_n(s),y)||z|Q_n(d\mathbf{v})\\ \leq & \int_{[t_0,t_1]\times\mathbb{L}\times\mathcal{M}_F\times\mathbb{R}^d}\kappa_1(|\xi_n(s)|+1)Q_n(d\mathbf{v})\\ & +\sqrt{\int_{[t_0,t_1]\times\mathbb{L}\times\mathcal{M}_F\times\mathbb{R}^d}|a(\xi_n(s),y)|^2Q_n(d\mathbf{v})\int_{\mathbb{H}_T}|z|^2Q_n(d\mathbf{v})}\\ \leq & \kappa_1(\sqrt{M_1}+1)|t_1-t_0|+2\kappa_1\sqrt{(M+1)(M_1+1)}|t_1-t_0|^{1/2}.\end{aligned}$$

Thus

$$\lim_{\delta\to 0}\sup_{n\geq 1}\sup_{|t_1-t_0|\leq\delta}|\xi_n(t_1)-\xi_n(t_0)|=0.$$

Pre-compactness of $\{\xi_n\}$ in $\mathbb{U}$ follows, and thus the proof of (i) is complete.

Proof of (ii). Let (Q,ξ) be a limit point of the sequence $\{(Q_n,\xi_n)\}_{n\geq 1}$. Part (a) is an immediate consequence of the lower semi-continuity of the map

$$Q\mapsto\int_{\mathbb{H}_T}\left[\frac{1}{2}|z|^2+\hat{l}(\eta)\right]Q(d\mathbf{v}).$$

Assume without loss of generality that the full sequence converges to (Q,ξ). For Part (b), note that from Lipschitz property of a and b,

$$\int_{\mathbb{H}_T}|b(\xi_n(s),y)-b(\xi(s),y)|Q_n(d\mathbf{v})+\int_{\mathbb{H}_T}|a(\xi_n(s),y)-a(\xi(s),y)||z|Q_n(d\mathbf{v})\to 0,$$

as $n\to\infty$. Also $(s,y,\eta,z)\mapsto b(\xi(s),y)$ is a continuous and bounded map, and $(s,y,\eta,z)\mapsto a(\xi(s),y)z$ is a continuous map and $\int_{\mathbb{H}_T}|z|^2Q_n(d\mathbf{v})\leq 2(M+1)$, from which it follows that

$$\int_{\mathbb{H}_T}[b(\xi(s),y)+a(\xi(s),y)z]Q_n(d\mathbf{v})\to\int_{\mathbb{H}_T}[b(\xi(s),y)+a(\xi(s),y)z]Q(d\mathbf{v}).$$

Combining the above two convergence statements we have (b).

Next we consider part (c). We will use the following inequality: For $u,v\in(0,\infty)$ and $\sigma\in[1,\infty)$

$$uv\leq e^{\sigma u}+\frac{1}{\sigma}(v\log v-v+1)=e^{\sigma u}+\frac{1}{\sigma}l(v). \tag{5.3.4}$$

A simple calculation using (5.2.1) and the above inequality shows that there exists $c_1\in(0,\infty)$ such that for all $x,x'\in\mathbb{R}^d$, $\eta\in\mathcal{M}_F$ and $m\in(1,\infty)$,

$$\sup_{y,y'\in\mathbb{L}}\left\{|\hat{c}^\eta(x,y)-\hat{c}^\eta(x',y)|+|\hat{c}^\eta(x,y)\hat{R}^\eta(x,y)-\hat{c}^\eta(x',y)\hat{R}^\eta(x',y)|\right\}$$
$$\leq c_1(e^m\kappa_2|x-x'|+\frac{\hat{l}(\eta)}{m}) \tag{5.3.5}$$

$$\sup_{y\in\mathbb{L}}|\Pi^\eta_x\phi(y)-\Pi^\eta_{x'}\phi(y)|\leq c_1|\phi|_\infty(e^m\kappa_2|x-x'|+\frac{\hat{l}(\eta)}{m}). \tag{5.3.6}$$

For completeness, we include the proof of above inequalities in the Appendix. From this estimate, along with the observation that $\int_{\mathbb{H}_T} \hat{l}(\eta) Q_n(d\mathbf{v}) \leq M+1$, we obtain that

$$\int_{\mathbb{H}_T} \left(\Pi^\eta_{\xi_n(s)} \phi(y) - \Pi^\eta_{\xi(s)} \phi(y) \right) Q_n(d\mathbf{v}) \to 0, \tag{5.3.7}$$

as $n \to \infty$.

Finally, using Lemma 5.3.1 below, an application of Skorohod representation theorem shows that

$$\int_{\mathbb{H}_T} \Pi^\eta_{\xi(s)} \phi(y) Q_n(d\mathbf{v}) \to \int_{\mathbb{H}_T} \Pi^\eta_{\xi(s)} \phi(y) Q(d\mathbf{v}). \tag{5.3.8}$$

Combining this with (5.3.7), we have (c). The result follows. □

Lemma 5.3.1. *Let (η^n, Z^n, Y^n) be a sequence of $\mathcal{M}_F \times \mathbb{R}^d \times \mathbb{L}$ valued random variables given on a probability space $(\bar{\Omega}, \bar{\mathcal{F}}, \bar{\mathbb{P}})$, which converges a.s. to $(\bar{\eta}, \bar{Z}, \bar{Y})$. Further suppose that, for some $C_1 \in (0, \infty)$, $\sup_{n \geq 1} \bar{\mathbb{E}} \hat{l}(\eta^n) \leq C_1$. Then for all $\phi \in \mathbb{M}(\mathbb{L})$, $\Pi^{\eta^n}_{Z^n} \phi(Y^n)$ converges a.s. to $\Pi^{\bar{\eta}}_{\bar{Z}} \phi(\bar{Y})$, as $n \to \infty$.*

Proof. Let $\Omega_0 \subset \bar{\Omega}$ be such that $\bar{\mathbb{P}}(\Omega_0) = 1$ and $\forall \omega \in \Omega_0$,

$$(\eta^n(\omega), Z^n(\omega), Y^n(\omega)) \to (\bar{\eta}(\omega), \bar{Z}(\omega), \bar{Y}(\omega)),$$

and $\hat{l}(\bar{\eta}(\omega)) < \infty$.

Using (5.3.6) we see that

$$|\Pi^{\eta^n}_{Z^n} \phi(Y^n) - \Pi^{\eta^n}_{\bar{Z}} \phi(Y^n)| \to 0, \quad \text{a.s.} \tag{5.3.9}$$

Also, note that for $(\eta, Z, Y) \in \mathcal{M}_F \times \mathbb{R}^d \times \mathbb{L}$,

$$\Pi^\eta_z \phi(y) = -\eta\{E_{z,y}\} \phi(y) + \sum_{y' \in \mathbb{L} \setminus y} \phi(y') \eta\{E^{y'}_{z,y}\} \tag{5.3.10}$$

Fix $\omega \in \Omega_0$. Then there exists a $N \equiv N(\omega)$ such that, $\forall n \geq N(\omega)$, $Y^n(\omega) = \bar{Y}(\omega)$, and thus (suppressing ω)

$$|\Pi_{\bar{Z}}^{\eta^n} \phi(Y^n) - \Pi_{\bar{Z}}^{\bar{\eta}} \phi(\bar{Y})| = |\Pi_{\bar{Z}}^{\eta^n} \phi(\bar{Y}) - \Pi_{\bar{Z}}^{\bar{\eta}} \phi(\bar{Y})|. \tag{5.3.11}$$

Recall from (5.2.2) that $\theta(\partial E_{z,y}) = 0$. Since $\hat{l}(\bar{\eta}(\omega)) < \infty$, we have that $\bar{\eta}(\partial E_{z,y}) = 0$. Combining this with $\eta^n \to \bar{\eta}$, we have that

$$\eta^n\{E_{z,y}\} \to \bar{\eta}\{E_{z,y}\}, \quad \forall (z, y) \in \mathbb{R}^d \times \mathbb{L}.$$

Similarly

$$\eta^n\{E_{z,y}^{y'}\} \to \bar{\eta}\{E_{z,y}^{y'}\}, \quad \forall (z, y, y') \in \mathbb{R}^d \times \mathbb{L} \times \mathbb{L}.$$

Using these observations in (5.3.10) and (5.3.11), we now have that

$$\Pi_{\bar{Z}}^{\eta^n} \phi(Y^n) \to \Pi_{\bar{Z}}^{\bar{\eta}} \phi(\bar{Y}), \qquad \text{a.s.}$$

Combining this with (5.3.9) we have the result. $\square$

5.4 Large Deviation Upper Bound.

In this section we will show that for all $F \in C_b(\mathbb{U})$,

$$\underline{\lim}_{\epsilon \to 0} -\epsilon \log \mathbb{E}_x \left[\exp \left(-\frac{1}{\epsilon} F(X^\epsilon) \right) \right] \geq \inf_{\xi \in \mathbb{U}} \{F(\xi) + \mathbb{I}(\xi)\}. \tag{5.4.1}$$

Let $\mathbb{M}$ denote the space of σ-finite measures on $[0,1] \times \mathbb{R}_+$ endowed with the vague topology. With this topology, for $\mu_n, \mu \in \mathbb{M}$, $\mu_n \to \mu$, if and only if $\int f(r,t)\mu_n(drdt) \to \int f(r,t)\mu(drdt)$ for all continuous real functions f on $[0,1] \times \mathbb{R}_+$ with compact support. Let $\mathbb{V} = \mathbb{U} \times \mathbb{M}$. Since there is a pathwise unique solution of (5.2.3)-(5.2.4), we have that, for each $\epsilon > 0$, there is a measurable map $\mathcal{G}^\epsilon : \mathbb{V} \to \mathbb{U}$ such that $X^\epsilon = \mathcal{G}^\epsilon(\sqrt{\epsilon} W, \epsilon N^{\epsilon^{-1}})$.

The proof of the upper bound relies on a variational representation from [18] which we now describe. Denote by $\mathcal{P}$ the predictable σ-field on $[0,T] \times \Omega$ with the filtration $\{\mathcal{F}_t : 0 \le t \le T\}$. Let $\bar{\mathcal{A}}$ be the class of all $(\mathcal{P}\otimes\mathcal{B}[0,1]\backslash\mathcal{B}[0,\infty))$ measurable maps $\varphi : [0,1]\times[0,T]\times\Omega \to [0,\infty)$.

Let

$$\mathcal{P}_2 = \left\{\psi = (\psi_i)_{i=1}^d : \psi_i \text{ is } \mathcal{P}\backslash\mathcal{B}(\mathbb{R}) \text{ measurable and } \int_0^T |\psi(s)|^2 ds < \infty, \text{ a.s. } \mathbb{P}\right\}$$

and set $\mathcal{U} = \mathcal{P}_2 \times \bar{\mathcal{A}}$. For $\psi \in \mathcal{P}_2$ define $\tilde{L}_T(\psi) = \frac{1}{2}\int_0^T |\psi(s)|^2 ds$ and for $\varphi \in \bar{\mathcal{A}}$, let $L_T(\varphi) = \int_{[0,1]\times[0,T]} l(\varphi(r,s))\theta(dr)ds$. For $\boldsymbol{u} = (\psi,\varphi) \in \mathcal{U}$, set $\bar{L}_T(\boldsymbol{u}) = L_T(\varphi) + \tilde{L}_T(\psi)$.

With this notation the variational representation of [18] says that

$$\begin{aligned}
-\epsilon \log \mathbb{E}_x &\left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right] \\
&= \inf_{\boldsymbol{u}=(\psi,\varphi)\in\mathcal{U}} \bar{\mathbb{E}}\left[\bar{L}_T(\boldsymbol{u}) + F\circ\mathcal{G}^\epsilon\left(\sqrt{\epsilon}W + \int_0^\cdot \psi(s)ds,\, \epsilon N^{\epsilon^{-1}\varphi}\right)\right].
\end{aligned} \tag{5.4.2}$$

In fact, a closer inspection of the proof of Theorem 2.8 of [18] shows that (5.4.2) can be strengthened as follows. For $n \ge 1$, define

$$\bar{\mathcal{A}}_{b,n} = \{\varphi \in \bar{\mathcal{A}} : \varphi(r,s,\omega) \in [n^{-1}, n], \text{ for all } (r,s,\omega) \in [0,1]\times[0,T]\times\Omega\}$$

and let $\bar{\mathcal{A}}_b = \cup_{n=1}^\infty \bar{\mathcal{A}}_{b,n}$. Also let $\mathcal{U}_b = \mathcal{P}_2 \times \bar{\mathcal{A}}_b$. Then in the equality in (5.4.2), $\mathcal{U}$ on the right side can be replaced by $\mathcal{U}_b$.

Let for $\epsilon > 0$, $\boldsymbol{u}^\epsilon = (\psi^\epsilon, \varphi^\epsilon) \in \mathcal{U}_b$ be such that

$$\begin{aligned}
-\epsilon \log \mathbb{E}_x &\left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right] \\
&\ge \bar{\mathbb{E}}\left[\bar{L}_T(\boldsymbol{u}^\epsilon) + F\circ\mathcal{G}^\epsilon\left(\sqrt{\epsilon}W + \int_0^\cdot \psi^\epsilon(s)ds,\, \epsilon N^{\epsilon^{-1}\varphi^\epsilon}\right)\right] - \epsilon.
\end{aligned} \tag{5.4.3}$$

Using the boundedness of F and a localization argument, one can assume without loss of generality that for some $M \in (0,\infty)$,

$$\sup_\epsilon \bar{L}_T(\boldsymbol{u}^\epsilon) \le M. \tag{5.4.4}$$

From unique pathwise solvability of (5.2.3) - (5.2.4), it follows that

$$\tilde{X}^\epsilon = \mathcal{G}^\epsilon\left(\sqrt{\epsilon}W + \int_0^\cdot \psi^\epsilon(s)ds,\, \epsilon N^{\epsilon^{-1}\varphi^\epsilon}\right) \text{ is the unique solution of (6.1.10)-(5.4.7) .} \tag{5.4.5}$$

For completeness, proof of (5.4.5) is included in the Appendix.

$$d\tilde{X}^\epsilon(t) = b(\tilde{X}^\epsilon(t), \tilde{Y}^\epsilon(t))dt + \sqrt{\epsilon}a(\tilde{X}^\epsilon(t), \tilde{Y}^\epsilon(t))dW(t) + a(\tilde{X}^\epsilon(t), \tilde{Y}^\epsilon(t))\psi^\epsilon(t)dt \tag{5.4.6}$$

$$d\tilde{Y}^\epsilon(t) = \int_{r\in[0,1]} k(\tilde{X}^\epsilon(t), \tilde{Y}^\epsilon(t), r)N^{\frac{1}{\epsilon}\varphi^\epsilon}(dr \times dt). \tag{5.4.7}$$

Using the linear growth property of b and a, and property (5.4.4), it is easy to check that

$$\sup_{\epsilon>0} \mathbb{E}\left(\sup_{0\le t\le T} |\tilde{X}^\epsilon(t)|^2\right) < \infty. \tag{5.4.8}$$

Similarly it can be verified that for some $c_1 \in (0, \infty)$,

$$\sup_{\epsilon>0}\sup_{0\le s\le T-\Delta} \mathbb{E}\left(\sup_{0\le t\le \Delta} |\tilde{X}^\epsilon(s+t) - \tilde{X}^\epsilon(s)|^2\right) \le c_1\Delta, \tag{5.4.9}$$

for any $\Delta \in [0, T]$. From (5.4.8) it follows that

$$\sup_{0\le t\le T} \epsilon\mathbb{E}\left|\int_0^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))dW(s)\right|^2 \to 0. \tag{5.4.10}$$

Fix $\{\Delta_\epsilon\}_{\epsilon>0}$ such that $\Delta_\epsilon \to 0$ and $\frac{\Delta_\epsilon}{\epsilon} \to \infty$ as $\epsilon \to 0$. Using (5.4.8) and (5.4.9) we have that

$$\sup_{0\le t\le T} \mathbb{E}\left|\int_0^t b(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))ds - \int_0^t \frac{1}{\Delta_\epsilon}\int_s^{(s+\Delta_\epsilon)\wedge T} b(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))duds\right|^2 \to 0, \tag{5.4.11}$$

$$\sup_{0\le t\le T} \mathbb{E}\left|\int_0^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))\psi^\epsilon(s)ds - \int_0^t \frac{1}{\Delta_\epsilon}\int_s^{(s+\Delta_\epsilon)\wedge T} a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)duds\right|^2 \to 0, \tag{5.4.12}$$

as $\epsilon \to 0$. Proof of (5.4.12) is given in the Appendix, and proof of (5.4.11) is very similar to that of (5.4.12), and thus omitted.

Adopting the convention that $(\tilde{Y}^\epsilon(u), \psi^\epsilon(u), \phi^\epsilon(r,u)) = (\tilde{Y}^\epsilon(T), 0, 1)$, for $u > T$, define $Q^\epsilon \in \mathcal{P}_1$ as

$$Q^\epsilon(A \times B \times C \times D) = \int_{[0,T]} 1_A(s) \left(\frac{1}{\Delta_\epsilon} \int_s^{s+\Delta_\epsilon} 1_B(\tilde{Y}^\epsilon(u)) 1_C(\eta^\epsilon(u)) 1_D(\psi^\epsilon(u)) du \right) ds,$$

where, for $u > 0$, $\eta^\epsilon(u) \in \mathcal{M}_F$ is defined as

$$\eta^\epsilon(u)(F) = \int_F \varphi^\epsilon(u,r)\theta(dr), \quad F \in \mathcal{B}[0,1]. \tag{5.4.13}$$

In terms of Q^ϵ, using (5.4.10) -(5.4.12) one can rewrite the evolution of $\tilde{X}^\epsilon$ as

$$\tilde{X}^\epsilon(t) = x + \int_{\mathbb{H}_t} b(\tilde{X}^\epsilon(s), y) Q^\epsilon(d\mathbf{v}) + \int_{\mathbb{H}_t} a(\tilde{X}^\epsilon(s), y) z Q^\epsilon(d\mathbf{v}) + \tilde{Z}^\epsilon(t), \tag{5.4.14}$$

where

$$\sup_{0 \le t \le T} \mathbb{E}|\tilde{Z}^\epsilon(t)|^2 \to 0, \tag{5.4.15}$$

as $\epsilon \to 0$.

Also the following inequality holds

$$\left[\bar{L}_T(u^\epsilon)\right] \ge \int_{\mathbb{H}_T} \left[\frac{1}{2}|z|^2 + \hat{l}(\eta)\right] Q^\epsilon(d\mathbf{v}), \tag{5.4.16}$$

Proof of (5.4.16) is given in the Appendix.

From (5.4.8) and using an estimate similar to (5.4.9) (in terms of stopping times) we have that $\{\tilde{X}^\epsilon\}_{\epsilon \ge 0}$ is a tight family of $\mathbb{U}$ valued random variables.

We will now prove the following statements:

(i) Q^ϵ is a tight family of $\mathcal{M}_F(\mathbb{H})$ valued random variables.

(ii) If $(\tilde{X}^0, Q^0)$ is a weak limit point of $(\tilde{X}^\epsilon, Q^\epsilon)$, then

(a) $\int_{\mathbb{H}_T} \left[\frac{1}{2}|z|^2 + \hat{l}(\eta)\right] Q^0(d\mathbf{v}) \le \liminf_{\epsilon \to 0} \bar{L}_T(u^\epsilon)$.

(b) Equation (5.2.9) holds with (ξ, Q) replaced by $(\tilde{X}^0, Q^0)$.

(c) (5.2.10) holds a.s. with (ξ, Q) replaced by $(\tilde{X}^0, Q^0)$.

Proof of (i). To show the tightness of $\{Q^\epsilon\}$, it is enough to show that for any $\delta > 0$, there exists $c(\delta)$ such that the following inequalities hold.

$$\sup_\epsilon \mathbb{E}Q^\epsilon\{(s,y,\eta,z) \in \mathbb{H}_T | \eta[0,1] > c(\delta)\} \leq \delta;$$

$$\sup_\epsilon \mathbb{E}Q^\epsilon\{(s,y,\eta,z) \in \mathbb{H}_T | \ |z| > c(\delta)\} \leq \delta.$$

Both of these inequalities are a consequence of (5.4.4) and (5.4.16); the proof of the first inequality follows from similar argument as in the proof of (5.3.2) while the proof of the second inequality is immediate on using Markov's inequality.

Proof of (ii). Part (a) follows as in the proof of Proposition 5.3.1. Consider now part (b). Assume without loss of generality that $(\tilde{X}^\epsilon, Q^\epsilon) \to (\tilde{X}^0, Q^0)$ a.s. Then, once again, as in the proof of Proposition 5.3.1, we have that

$$\begin{aligned}\int_{\mathbb{H}_T} &[b(\tilde{X}^\epsilon(s), y) + a(\tilde{X}^\epsilon(s), y)z]Q^\epsilon(d\mathbf{v}) \\ &\to \int_{\mathbb{H}_T} [b(\tilde{X}^0(s), y) + a(\tilde{X}^0(s), y)z]Q^0(d\mathbf{v}).\end{aligned}$$

Combining this with (5.4.14) and (5.4.15), we have part (b).

For part (c), we first estimate the difference between $\int_{\mathbb{H}_t} \Pi^\eta_{\tilde{X}^\epsilon(s)}\phi(y)Q^\epsilon(d\mathbf{v})$ and $\int_0^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))du$. By a change of the order of integration, we have

$$\begin{aligned}\int_{\mathbb{H}_t} \Pi^\eta_{\tilde{X}^\epsilon(s)}\phi(y)Q^\epsilon(d\mathbf{v}) &= \int_0^t \frac{1}{\Delta_\epsilon}\int_s^{s+\Delta_\epsilon} \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u))duds \\ &= \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_0^u \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u))dsdu \\ &\quad + \int_{\Delta_\epsilon}^t \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u))dsdu \\ &\quad + \int_t^{t+\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u))dsdu.\end{aligned}$$

Also, we have

$$\int_0^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))du = \int_0^t \frac{1}{(\Delta_\epsilon)\wedge u}\int_{(u-\Delta_\epsilon)_+}^u \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))dsdu.$$

Using (5.4.4) and (5.3.5), there exists $c \in (0,\infty)$ such that, for any $m \in [1,\infty)$

$$\begin{aligned}
&\left|\int_{\mathbb{H}_t} \Pi^{\eta}_{\tilde{X}^\epsilon(s)}\ \phi(y)Q^\epsilon(d\mathbf{v}) - \int_0^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))du\right| \\
&\qquad\qquad \leq \left|\int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_0^u \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u))dsdu\right| + \left|\int_0^{\Delta_\epsilon} \frac{1}{u}\int_0^u \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))dsdu\right| \\
&\qquad\qquad\quad + \int_{\Delta_\epsilon}^t \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u \left|\Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u)) - \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))\right| dsdu \\
&\qquad\qquad\quad + \left|\int_t^{t+\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(s)}\phi(\tilde{Y}^\epsilon(u))dsdu\right| \\
&\qquad\qquad \leq ce^m|\Delta_\epsilon| + ce^m \sup_{|s-s'|\leq\Delta_\epsilon, s,s'\in[0,T]} |\tilde{X}^\epsilon(s) - \tilde{X}^\epsilon(s')| + \frac{c}{m}\int_0^T \hat{l}(\eta^\epsilon(u))du
\end{aligned}$$

for all $t \in [0,T]$. Thus

$$\begin{aligned}
&\limsup_{\epsilon\to 0}\left|\int_{\mathbb{H}_t} \Pi^{\eta}_{\tilde{X}^\epsilon(s)}\phi(y)Q^\epsilon(d\mathbf{v}) - \int_0^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))du\right| \\
&\leq \limsup_{m\to\infty}\left[ce^m \limsup_{\epsilon\to 0}\sup_{|s-s'|\leq\Delta_\epsilon, s,s'\in[0,T]}\sup_{\tilde{\epsilon}>0}|\tilde{X}^{\tilde{\epsilon}}(s) - \tilde{X}^{\tilde{\epsilon}}(s')| + \frac{c}{m}\sup_{\tilde{\epsilon}>0}\int_0^T \hat{l}(\eta^{\tilde{\epsilon}}(u))du\right] \\
&\leq \limsup_{m\to\infty}\frac{c}{m}\sup_{\tilde{\epsilon}>0}\int_0^T \hat{l}(\eta^{\tilde{\epsilon}}(u))du \\
&= 0 \qquad\qquad (5.4.17)
\end{aligned}$$

where the second inequality follows on noting that $\{\tilde{X}^\epsilon, \epsilon > 0\}$ is an equicontinuous family and the final equality follows from (5.4.4). Also, using Itô's formula, we have

$$\begin{aligned}
&\phi(\tilde{Y}^\epsilon(t)) - \phi(y_0) \\
&= \int_{[0,1]\times[0,t]} \phi(\tilde{Y}^\epsilon(u-) + k(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u-), r)) - \phi(\tilde{Y}^\epsilon(u-))N^{\frac{1}{\epsilon}\varphi^\epsilon}(drdu) \\
&= \int_{[0,1]\times[0,t]} \phi(\tilde{Y}^\epsilon(u-) + k(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u-), r)) - \phi(\tilde{Y}^\epsilon(u-))\tilde{N}^{\frac{1}{\epsilon}\varphi^\epsilon}(drdu) \\
&\quad + \int_{[0,1]\times[0,t]} \phi(\tilde{Y}^\epsilon(u) + k(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u), r)) - \phi(\tilde{Y}^\epsilon(u))\frac{1}{\epsilon}\varphi^\epsilon(r,u)\theta(dr)du \\
&= \int_{[0,1]\times[0,t]} \phi(\tilde{Y}^\epsilon(u-) + k(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u-), r)) - \phi(\tilde{Y}^\epsilon(u-))\tilde{N}^{\frac{1}{\epsilon}\varphi^\epsilon}(drdu) \\
&\quad + \frac{1}{\epsilon}\int_0^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)}\phi(\tilde{Y}^\epsilon(u))du.
\end{aligned}$$

Rearranging terms,

$$
\begin{aligned}
&\int_0^t \Pi^{\eta^\epsilon(u)}_{\tilde{X}^\epsilon(u)} \phi(\tilde{Y}^\epsilon(u))du \\
=&\epsilon \left(\phi(\tilde{Y}^\epsilon(t)) - \phi(y_0) \right) \\
&- \epsilon \int_{[0,1]\times[0,t]} \left[\phi(\tilde{Y}^\epsilon(u-) + k(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u), r)) - \phi(\tilde{Y}^\epsilon(u-)) \right] \tilde{N}^{\frac{1}{\epsilon}\varphi^\epsilon}(drdu),
\end{aligned}
$$

which clearly converges to 0 in probability, as $\epsilon \to 0$. Combining this observation with (5.4.17), we have

$$
\int_{\mathbb{H}_t} \Pi^{\eta}_{\tilde{X}^\epsilon(s)} \phi(y) Q^\epsilon(d\mathbf{v}) \to 0 \quad \text{ in probability.}
$$

Finally using the estimates in (5.3.6) and Lemma 5.3.1 once more, we have

$$
\int_{\mathbb{H}_t} \Pi^{\eta}_{\tilde{X}^0(s)} \phi(y) Q^0(d\mathbf{v}) = 0 \quad \text{a.s.},
$$

which completes the proof of part (c) and hence that of (ii).

From (i) and (ii) we now have that $\{Q^\epsilon\}$ is tight and if $(Q^0, \tilde{X}^0)$ is a limit point of $(Q^\epsilon, \tilde{X}^\epsilon)$, then $Q^0 \in \mathcal{A}_{\tilde{X}^0}$ a.s.. Taking limit as $\epsilon \to 0$ (along the subsequence) in (5.4.3), we now see that

$$
\begin{aligned}
\liminf_{\epsilon\to 0} -\epsilon \log \mathbb{E}_x &\left[\exp\left(-\frac{1}{\epsilon} F(X^\epsilon) \right) \right] \\
&\geq \mathbb{E}\left[\left(\int_{\mathbb{H}_{\mathbb{T}}} \left[\frac{1}{2}|z|^2 + \hat{l}(\eta) \right] Q^0(d\mathbf{v}) \right) + F(\tilde{X}^0) \right] \\
&\geq \mathbb{E}\left[\inf_{Q \in \mathcal{A}_{\tilde{X}^0}} \left(\int_{\mathbb{H}_{\mathbb{T}}} \left[\frac{1}{2}|z|^2 + \hat{l}(\eta) \right] Q(d\mathbf{v}) \right) + F(\tilde{X}^0) \right] \\
&= \mathbb{E}\left[\mathbb{I}(\tilde{X}^0) + F(\tilde{X}^0) \right] \\
&\geq \inf_{\xi \in \mathbb{U}} \left[\mathbb{I}(\xi) + F(\xi) \right].
\end{aligned}
$$

This completes the proof of the upper bound. □

5.5 Large Deviation Lower Bound.

In this section we will show that for all $F \in C_b(\mathbb{U})$,

$$\overline{\lim_{\epsilon\to 0}} -\epsilon \log \mathbb{E}_x \left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right] \le \inf_{\xi\in\mathbb{U}}\{F(\xi)+\mathbb{I}(\xi)\} \tag{5.5.1}$$

where X^ϵ is given through (5.2.3)–(5.2.4) and $\mathbb{I}$ is as defined in (5.2.11). The section is organized as follows. In Section 5.5.1, we introduce the local rate function and a suitable regularization of the function that allows for certain approximation arguments. Next, in Section 5.5.2 we give some auxiliary lemmas for the proof of LDP lower bound. Finally, in Section 5.5.3 we complete the proof of the lower bound.

5.5.1 Local Rate Function.

In this section we will give an alternative expression for the rate function which is more amenable for the proof of the lower bound. Let

$$\mathcal{K} = \{q : \mathbb{L}\times\mathcal{B}(\mathcal{M}_F) \to [0,1] \mid q(y,\cdot)\in\mathcal{P}^*(\mathcal{M}_F),\ \text{for all } y\in\mathbb{L}\},$$

where $\mathcal{P}^*(\mathcal{M}_F)$ is the space of all probability measures μ on $\mathcal{M}_F$ with $\int \hat{l}(\eta)\mu(d\eta) < \infty$. For $q\in\mathcal{K}$, let $\vartheta^q : \mathbb{L}\times\mathcal{B}[0,1]\to[0,\infty]$ be defined as

$$\vartheta^q(y,A) = \int_{\mathcal{M}_F} \eta(A)q(y,d\eta),\ (y,A)\in\mathbb{L}\times\mathcal{B}[0,1].$$

Then from superlinearity of l it follows that, for each $y\in\mathbb{L}$, $\vartheta^q(y,\cdot)$ is a finite measure on $[0,1]$. Define for q as above, $\hat{c}^q : \mathbb{R}^d\times\mathbb{L}\to[0,\infty)$ and $\hat{R}^q : \mathbb{R}^d\times\mathbb{L}\times\mathbb{L}\to[0,1]$ as

$$\begin{aligned}
\hat{c}^q(x,y) &= \int_{[0,1]} 1_{\{r:k(x,y,r)\neq 0\}}\vartheta^q(y,dr),\\
\hat{c}^q(x,y)\hat{R}^q(x,y,y') &= \int_{[0,1]} 1_{\{r:y+k(x,y,r)=y'\}}\vartheta^q(y,dr),\ (x,y)\in\mathbb{G},\ y'\in\mathbb{L}\setminus\{y\}.
\end{aligned}$$

We set $\hat{R}^q(x,y,y)=0$ for all $(x,y)\in\mathbb{G}$. Also, if $\hat{c}^q(x,y)=0$, by convention we take $\hat{R}^q(x,y,y')=R(x,y,y')$. Define $\hat{R}^q(x,y,A)=\sum_{y\in A}\hat{R}^q(x,y,y')$ for all $A\subset\mathbb{L}$.

For $q \in \mathcal{K}$, $(x,y) \in \mathbb{G}$ and $\phi \in \mathbb{M}(\mathbb{L})$, define

$$\begin{aligned} \hat{\Pi}_x^q \phi(y) &= -\,\hat{c}^q(x,y)\phi(y) + \hat{c}^q(x,y)\int_{\mathbb{L}} \phi(z)\hat{R}^q(x,y,dz) && (5.5.2)\\ &= \int \Pi_x^\eta \phi(y) q(y,d\eta), && (5.5.3) \end{aligned}$$

where Π_x^η is defined as in (5.2.7).

For $\varpi \in \mathcal{P}(\mathbb{L})$, $x \in \mathbb{R}^d$, let

$$\mathcal{K}(\varpi,x) = \left\{ q \in \mathcal{K} \,\middle|\, \int_{\mathbb{L}} \hat{\Pi}_x^q \phi(y)\varpi(dy) = 0, \quad \text{for all } \phi \in \mathbb{M}(\mathbb{L}) \right\}$$

Let $G_1 : \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d \to [0,\infty]$ be defined as

$$G_1(\varpi,x) = \inf_{q\in\mathcal{K}(\varpi,x)} \int_{\mathbb{L}} \int_{\mathcal{M}_F} \hat{l}(\eta) q(y,d\eta)\varpi(dy).$$

Next,for $(\varpi,x,v) \in \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d \times \mathbb{R}^d$, let

$$\mathcal{U}(\varpi,x,v) = \left\{ u : \mathbb{L} \to \mathbb{R}^d \,\middle|\, \int_{\mathbb{L}} \left(a(x,y)u(y) + b(x,y)\right)\varpi(dy) = v \right\}$$

and let $G_2 : \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d \times \mathbb{R}^d \to [0,\infty]$ be defined as

$$G_2(\varpi,x,v) = \inf_{u\in\mathcal{U}(\varpi,x,v)} \left\{ \frac{1}{2}\int_{\mathbb{L}} |u(y)|^2 \varpi(dy) \right\}.$$

Let $G(\varpi,x,v) = G_1(\varpi,x) + G_2(\varpi,x,v)$, and

$$L(x,v) = \inf_{\varpi\in\mathcal{P}(\mathbb{L})} \{G(\varpi,x,v)\},\ (x,v) \in \mathbb{R}^d \times \mathbb{R}^d. \tag{5.5.4}$$

It will be convenient to work with a somewhat different representation for G_1. Let

$$\bar{\mathcal{K}} = \{\vartheta : \mathbb{L} \times \mathcal{B}[0,1] \to [0,\infty) : \vartheta(y,\cdot) \in \mathcal{M}_F \text{ for all } y \in \mathbb{L}\}. \tag{5.5.5}$$

Given $\varpi \in \mathcal{P}(\mathbb{L})$ and $x \in \mathbb{R}^d$, let

$$\bar{\mathcal{K}}(\varpi,x) = \{\vartheta \in \bar{\mathcal{K}} : \int_{\mathbb{L}} \Pi_x^\vartheta \phi(y)\varpi(dy) = 0,\ \text{for all } \phi \in \mathbb{M}(\mathbb{L})\},$$

where, for $\vartheta \in \bar{\mathcal{K}}$ and $x \in \mathbb{R}^d$, Π_x^ϑ is the generator of a $\mathbb{L}$ valued Markov process defined as

$$\Pi_x^\vartheta \phi(y) = -\hat{c}^{\vartheta(y,\cdot)}(x,y)\phi(y) + \hat{c}^{\vartheta(y,\cdot)}(x,y) \int_{\mathbb{L}} \phi(z) \hat{R}^{\vartheta(y,\cdot)}(x,y,dz).$$

Let $\bar{G}_1 : \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d \to [0,\infty]$ be defined as

$$\bar{G}_1(\varpi, x) = \inf_{\vartheta \in \bar{\mathcal{K}}(\varpi,x)} \int_{\mathbb{L}} \hat{l}(\vartheta(y,\cdot))\varpi(dy).$$

The following lemma shows that G_1 and $\bar{G}_1$ are the same.

Lemma 5.5.1. *For all $(\varpi, x) \in \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d$, $\bar{G}_1(\varpi, x) = G_1(\varpi, x)$.*

Proof. For any $\vartheta \in \bar{\mathcal{K}}(\varpi, x)$, $q(y, d\eta) = \delta_{\vartheta(y.\cdot)}(d\eta)$ defines an element of $\mathcal{K}(\varpi, x)$ and clearly,

$$\int_{\mathbb{L}} \hat{l}(\vartheta(y,\cdot))\varpi(dy) = \int_{\mathbb{L}} \int_{\mathcal{M}_F} \hat{l}(\eta) q(y, d\eta)\varpi(dy).$$

Thus $\bar{G}_1(\varpi, x) \geq G_1(\varpi, x)$. Conversely, given a $q \in \mathcal{K}(\varpi, x)$, define $\vartheta^q \in \bar{\mathcal{K}}(\varpi, x)$ as

$$\vartheta^q(y, A) = \begin{cases} \int \eta(A) q(y, d\eta) & \text{if } \varpi(y) > 0, \\ \theta(A) & \text{otherwise.} \end{cases} \tag{5.5.6}$$

Without loss of generality we can assume that $\int_{\mathbb{L}} \int_{\mathcal{M}_F} \hat{l}(\eta) q(y, d\eta)\varpi(dy) < \infty$. Then, for any y with $\varpi(y) > 0$, we must have $\int \hat{l}(\eta) q(y, d\eta) < \infty$ and consequently $\vartheta^q(y,\cdot) \in \mathcal{M}_F$ for all $y \in \mathbb{L}$. It is immediate that $\vartheta^q \in \bar{\mathcal{K}}(\varpi, x)$. Also, using convexity of $\hat{l}$ and applying Jensen's inequality, we see

$$\bar{G}_1(\varpi, x) \leq \int_{\mathbb{L}} \hat{l}(\vartheta^q(y,\cdot))\varpi(dy) \leq \int_{\mathbb{L}} \int_{\mathcal{M}_F} \hat{l}(\eta) q(y, d\eta)\varpi(dy).$$

Since $q \in \mathcal{K}(\varpi, x)$ is arbitrary, the result follows. $\square$

Define, for $\varphi \in \mathbb{U}$,

$$\tilde{\mathbb{I}}(\varphi) = \begin{cases} \int_0^T L(\varphi_s, \dot{\varphi}_s)ds & \text{if } \int_0^T |\dot{\varphi}_s|^2 ds < \infty, \\ \infty & \text{otherwise.} \end{cases} \tag{5.5.7}$$

The following proposition allows us to work with $\tilde{\mathbb{I}}$, instead of $\mathbb{I}$, when proving the lower bound.

Proposition 5.5.1. *For all $\varphi \in \mathbb{U}$, $\mathbb{I}(\varphi) \geq \tilde{\mathbb{I}}(\varphi)$.*

Proof. Fix $\varphi \in \mathbb{U}$ with $\mathbb{I}(\varphi) < \infty$. For fixed $\epsilon_0 > 0$, choose $Q \in \mathcal{A}_\varphi$ such that

$$\mathbb{I}(\varphi) \geq \int_{\mathbb{H}_T} \left(\frac{1}{2}|z|^2 + \hat{l}(\eta)\right) Q(d\mathbf{v}) - \epsilon_0.$$

Disintegrate Q as

$$Q(ds\, dy\, d\eta\, dz) = Q_s(dy\, d\eta\, dz)ds.$$

Also, disintegrate the marginals

$$Q_s^{1,2}(dy\, d\eta) = \hat{Q}_s^2(y,\, d\eta)Q_s^1(dy),\;\; Q_s^{1,3}(dy\, dz) = \hat{Q}_s^3(y,\, dz)Q_s^1(dy).$$

Define $\vartheta_s(y, A) = \int_{\mathcal{M}_F} \eta(A)\hat{Q}_s^2(y,\, d\eta)$. It is easily checked that $\int_{\mathbb{L}} \Pi_{\varphi(s)}^{\vartheta_s(y,\cdot)} \phi(y) Q_s^1(dy) = 0$, for all $\phi \in \mathbb{M}(\mathbb{L})$, a.e. $s \in [0, T]$. Therefore

$$\vartheta_s \in \bar{\mathcal{K}}(Q_s^1, \varphi(s)), \quad \text{for a.e. } s \in [0, T]. \tag{5.5.8}$$

Also, let

$$\psi(s) = \int_{\mathbb{L}} \left(b(\varphi(s), y) + a(\varphi(s), y) \int_{\mathbb{R}^d} z\hat{Q}_s^3(y, dz)\right) Q_s^1(dy).$$

Then $\psi(s) = \dot{\varphi}(s)$ for a.e. $s \in [0, T]$ and using (5.2.8) it follows that

$$\int_0^T |\dot{\varphi}(s)|^2 ds = \int_0^T |\psi(s)|^2 ds < \infty.$$

Define $u_s(y) = \int_{\mathbb{R}^d} z\hat{Q}_s^3(y, dz)$. Then $\int |u_s(y)|^2 Q_s^1(dy) < \infty$ for a.e. $s \in [0, T]$, consequently

$$u_s \in \mathcal{U}(Q_s^1, \varphi(s), \psi(s)), \quad \text{for a.e. } s \in [0, T].$$

Also,

$$\begin{aligned}\int_{[0,T]} G_2(Q_s^1, \varphi(s), \psi(s))ds \leq & \frac{1}{2}\int_{[0,T]}\int_{\mathbb{L}} |u_s(y)|^2 Q_s^1(dy)ds \\ \leq & \int_{[0,T]\times\mathbb{L}\times\mathbb{R}^d} \frac{1}{2}|z|^2 \hat{Q}_s^3(y,dz)Q_s^1(dy)ds \\ = & \int_{\mathbb{H}_T} \frac{1}{2}|z|^2 Q(d\mathbf{v}).\end{aligned}$$

Similarly, using (5.5.8),

$$\begin{aligned}\int_{[0,T]} G_1(Q_s^1, \varphi(s))ds \leq & \int_{[0,T]}\int_{\mathbb{L}} \hat{l}(\vartheta_s(y,\cdot))Q_s^1(dy)ds \\ \leq & \int_{[0,T]\times\mathbb{L}\times\mathcal{M}_F} \hat{l}(\eta)\hat{Q}_s^2(y,d\eta)Q_s^1(dy)ds \\ = & \int_{\mathbb{H}_T} \hat{l}(\eta)Q(d\mathbf{v})\end{aligned}$$

Combining these estimates we have

$$\begin{aligned}\tilde{\mathbb{I}}(\varphi) = \int_{[0,T]} L(\varphi(s), \psi(s))ds \leq & \int_{[0,T]} G(Q_s^1, \varphi(s), \psi(s))ds \\ \leq & \int_{\mathbb{H}_T} (\frac{1}{2}|z|^2 + \hat{l}(\eta))Q(d\mathbf{v}) \leq \mathbb{I}(\varphi) + \epsilon_0.\end{aligned}$$

The desired inequality follows on sending ϵ_0 to 0. ☐

We will now introduce a regularization of L which will allow us to do certain approximation arguments. For $\sigma \in (0, \infty)$, define

$$G_{2,\sigma} : \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d \times \mathbb{R}^d \to [0, \infty]$$

as

$$G_{2,\sigma}(\varpi, x, v) = \inf_{r\in\mathbb{R}^d} \left[G_2(\varpi, x, v - r) + \frac{|r|^2}{2\sigma^2} \right]$$

and let

$$G_\sigma(\varpi, x, v) = G_1(\varpi, x) + G_{2,\sigma}(\varpi, x, v),\ (\varpi, x, v) \in \mathcal{P}(\mathbb{L}) \times \mathbb{R}^d \times \mathbb{R}^d.$$

Finally, let

$$L_\sigma(x, v) = \inf_{\varpi\in\mathcal{P}(\mathbb{L})} \{G_\sigma(\varpi, x, v)\},\ (x, v) \in \mathbb{R}^d \times \mathbb{R}^d.$$

Observe that

$$G_{2,\sigma}(\varpi, x, v) \le G_2(\varpi, x, v) \text{ and } L_\sigma(x, v) \le L(x, v), \text{ for all } (\varpi, x, v) \in \mathcal{P}(\mathbb{L})\times\mathbb{R}^d\times\mathbb{R}^d.$$

The following elementary lemma shows that L is locally bounded.

Lemma 5.5.2. *For every $M \in (0, \infty)$, there exists a $M_0 \in (0, \infty)$ such that*

$$L_\sigma(x, v) \le \frac{M_0}{\sigma^2}(1 + |v|^2), \text{ for all } \sigma > 0 \text{ and } (x, v) \in \mathbb{R}^d \times \mathbb{R}^d \text{ with } |x| \le M.$$

Proof. Note that for $\sigma \in (0, \infty)$ and $(x, v) \in \mathbb{R}^d \times \mathbb{R}^d$, $L_\sigma(x, v) \le G_\sigma(\rho_x, x, v)$ where ρ_x is as introduced in Theorem 5.2.1. Noting that $q(y, d\eta) = \delta_{\{\theta\}}(d\eta) \in \mathcal{K}(\rho_x, x)$, we see that $G_1(\rho_x, x) = 0$. Let $r = v - \int_{\mathbb{L}} b(x, y)\rho_x(dy)$. Then there exists $M_0 \in (0, \infty)$ such that $|r|^2 \le 2M_0(1+|v|^2)$ for all $x \in \mathbb{R}^d$ with $|x| \le M$. Also, clearly $u = 0$ belongs to $\mathcal{U}(\rho_x, x, v - r)$. Thus

$$L_\sigma(x, v) \le G_\sigma(\rho_x, x, v) = G_{2,\sigma}(\rho_x, x, v) \le G_2(\rho_x, x, v - r) + \frac{|r|^2}{2\sigma^2} \le \frac{2M_0(1 + |v|^2)}{2\sigma^2}.$$

The result follows. □

The following lemma is a consequence of an elementary Lagrange multiplier argument.

Lemma 5.5.3. *Let $K_1 \subset \mathbb{R}^d$ be a compact set. Then there exists a $B \in (0, \infty)$ such that for every $\sigma \in (0, \infty)$, $x \in K_1$, $v \in \mathbb{R}^d$ and $\varpi \in \mathcal{P}(\mathbb{L})$ there is a $r \in \mathbb{R}^d$ and $u \in \mathcal{U}(\varpi, x, v - r)$, with*

$$G_{2,\sigma}(\varpi, x, v) = \frac{1}{2}\int_{\mathbb{L}} |u(y)|^2\varpi(dy) + \frac{|r|^2}{2\sigma^2},$$

$|r| \le B(1 + |v|)$ and $|u(i)| \le \frac{B(1+|v|)}{\sigma^2}$ for all $i \in \mathbb{L}$.

Proof. For fixed (ϖ, x, v) the equality in the above display holds for a u that minimizes $\frac{1}{2}\sum_{i\in\mathbb{L}} |u(i)|^2\varpi(i) + \frac{|r|^2}{2\sigma^2}$, subject to the constraint

$$\sum_i (b(x,i) + a(x,i)u(i))\varpi(i) + r = v. \tag{5.5.9}$$

We set $u(i) = 0$ if $\varpi(i) = 0$ so we can assume without loss of generality that $\varpi(i) > 0$ for all $i \in \mathbb{L}$. For simplicity assume first that all quantities are scaler valued. Then differentiating

$$\frac{1}{2}\sum_{i\in\mathbb{L}} |u(i)|^2\varpi(i) + \frac{|r|^2}{2\sigma^2} - \lambda(\sum_i (b(x,i) + a(x,i)u(i))\varpi(i) + r - v)$$

with respect to $u(i)$ and setting the derivative to 0 we get $u(i) = \lambda a(x,i)$. Also, differentiating with respect to r we see that $r = \lambda\sigma^2$. Using this in (5.5.9) we get

$$\lambda = \frac{v - \sum_i b(x,i)\varpi(i)}{\sum_i a^2(x,i)\varpi(i) + \sigma^2}$$

and so

$$u(i) = \frac{v - \sum_i b(x,i)\varpi(i)}{\sum_i a^2(x,i)\varpi(i) + \sigma^2} a(x,i),\ r = \frac{v - \sum_i b(x,i)\varpi(i)}{\sum_i a^2(x,i)\varpi(i) + \sigma^2}\sigma^2.$$

In the general vector valued case, a similar argument shows that, letting $\mathbf{a}(i) = (a(x,i))_{jk}$, one can take

$$u(i) = \mathbf{a}'(i)\mathbf{M}_\sigma^{-1}[v - \sum_i b(x,i)\varpi(i)],\ r = \sigma^2\mathbf{M}_\sigma^{-1}[v - \sum_i b(x,i)\varpi(i)],$$

$$\mathbf{M}_\sigma = \sum_i \varpi(i)\mathbf{A}(i) + \sigma^2 \mathrm{Id},\ \mathbf{A}(i) = \mathbf{a}(i)\mathbf{a}'(i).$$

Finally, the result follows on observing that a, b are bounded on $K_1 \times \mathbb{L}$ and for $\alpha \in \mathbb{R}^d$ $\alpha'\mathbf{M}_\sigma\alpha \geq \sigma^2\alpha'\alpha$. $\square$

Lemma 5.5.4. *Let $\epsilon_0 \in (0,1)$. Then there exists a $c_1 \in (0,\infty)$ such that for every $\varpi \in \mathcal{P}(\mathbb{L})$ and $x \in \mathbb{R}^d$, there is a $\vartheta \in \bar{\mathcal{K}}(\varpi, x)$ such that*

$$\int_{\mathbb{L}} \hat{l}(\vartheta(y,\cdot))\varpi(dy) \leq G_1(\varpi, x) + \epsilon_0$$

and $\varpi(y)\frac{d\vartheta(y,\cdot)}{d\theta} \leq c_1$ for every $y \in \mathbb{L}$.

Proof. Fix $\epsilon_0 \in (0,1)$. Choose $q_1 \in \mathcal{K}(\varpi, x)$ such that

$$G_1(\varpi, x) \geq \int_{\mathbb{L}} \int_{\mathcal{M}_F} \hat{l}(\eta) q_1(y, d\eta) \varpi(dy) - \epsilon_0.$$

Then $\vartheta^{q_1} \in \bar{\mathcal{K}}(\varpi, x)$, and

$$\int_{\mathbb{L}} \hat{l}(\vartheta^{q_1}(y, \cdot)) \varpi(dy) \leq G_1(\varpi, x) + \epsilon_0. \tag{5.5.10}$$

Now we will modify ϑ^{q_1} as follows. Define for $y \in \mathbb{L}$, the measure $\tilde{\vartheta}(y, \cdot)$ as

$$\frac{d\tilde{\vartheta}(y, \cdot)}{d\theta}(r) = \begin{cases} \alpha_{x,y}^{y'} & \text{if } r \in E_{x,y}^{y'} \text{ for some } y' \in \mathbb{L} \text{ and } \theta(E_{x,y}^{y'}) \neq 0, \\ 1 & \text{otherwise.} \end{cases}$$

where $\alpha_{x,y}^{y'} = \frac{\vartheta^{q_1}(y, E_{x,y}^{y'})}{\theta(E_{x,y}^{y'})}$. From convexity of l and (5.5.10) it follows that

$$\begin{aligned} \int \hat{l}(\tilde{\vartheta}(y, \cdot)) \varpi(dy) &= \int \int_{[0,1]} l(\frac{d\tilde{\vartheta}(y, \cdot)}{d\theta}(r)) \theta(dr) \varpi(dy) \\ &\leq \int \hat{l}(\vartheta^{q_1}(y, \cdot)) \varpi(dy) \leq G_1(\varpi, x) + \epsilon_0. \end{aligned}$$

Also, $\Pi_x^{\vartheta^{q_1}} = \Pi_x^{\tilde{\vartheta}}$. Thus $\tilde{\vartheta} \in \overline{\mathcal{K}}(\varpi, x)$, and so does $\alpha\tilde{\vartheta}$, for any $\alpha \in (0, \infty)$. Denote

$$M = \sum_{y \in \mathbb{L}} \left(\sum_{y' \in \mathbb{L}/y} \tilde{\vartheta}(y, E_{x,y}^{y'}) + \theta(\overline{E}_{x,y}) \right) \varpi(y), \text{ and } c = \sum_{y \in \mathbb{L}} \theta(\overline{E}_{x,y}) \varpi(y),$$

where $\overline{E}_{x,y} = [0,1] \setminus E_{x,y}$ Define $\tilde{\vartheta}^*(y, \cdot) = \frac{\tilde{\vartheta}(y, \cdot)}{M}$. Note that $\sum_{y \in \mathbb{L}} \tilde{\vartheta}^*(y, E_{x,y}) \varpi(y) = \frac{M-c}{M}$. Also

$$\inf_{\alpha > 0} \int \hat{l}(\alpha \tilde{\vartheta}^*(y, \cdot)) \varpi(dy) \leq \int \hat{l}(\tilde{\vartheta}(y, \cdot)) \varpi(dy).$$

Next note that

$$\int \hat{l}(\alpha \tilde{\vartheta}^*(y, \cdot)) \varpi(dy) = \sum_{y \in \mathbb{L}} \sum_{y' \in \mathbb{L}/y} l(\alpha \frac{\tilde{\vartheta}^*(y, E_{x,y}^{y'})}{\theta(E_{x,y}^{y'})}) \theta(E_{x,y}^{y'}) \varpi(y) + l(\alpha/M) c.$$

Denote $a_y^{y'} = \tilde{\vartheta}^*(y, E_{x,y}^{y'})$, $b_y^{y'} = \theta(E_{x,y}^{y'})$. Taking derivative with respect to α, and setting it to 0, we have

$$\log(\alpha) [\sum_{y \in \mathbb{L}} \sum_{y' \in \mathbb{L}/y} a_y^{y'} \varpi(y) + \frac{c}{M}] = \sum_{y \in \mathbb{L}} \sum_{y' \in \mathbb{L}/y} a_y^{y'} (\log b_y^{y'} - \log a_y^{y'}) \varpi(y) + \frac{c \log M}{M}.$$

Solving for α, we have

$$\begin{aligned}\log(\alpha) &= \frac{\sum_{y\in\mathbb{L}}\sum_{y'\in\mathbb{L}\setminus y} a_y^{y'}(\log b_y^{y'} - \log a_y^{y'})\varpi(y) + \frac{c\log M}{M}}{\sum_{y\in\mathbb{L}}\sum_{y'\in\mathbb{L}\setminus y} a_y^{y'}\varpi(y) + \frac{c}{M}} \\ &= \sum_{y\in\mathbb{L}}\sum_{y'\in\mathbb{L}\setminus y} a_y^{y'}\log b_y^{y'}\varpi(y) - \sum_{y\in\mathbb{L}}\sum_{y'\in\mathbb{L}\setminus y} a_y^{y'}\log a_y^{y'}\varpi(y) + \frac{c\log M}{M},\end{aligned}$$

where the last equality follows on observing that the denominator in the first equality equals $\frac{M-c}{M} + \frac{c}{M} = 1$. Next note that

$$\begin{aligned}&\left(\frac{c\log M}{M}\right)_+ \leq c \leq \theta[0,1], \\ &\left(\sum_{y\in\mathbb{L}}\sum_{y'\in\mathbb{L}\setminus y} a_y^{y'}\log a_y^{y'}\varpi(y)\right)_- \leq \frac{\ell}{e}, \\ &\left(\sum_{y\in\mathbb{L}}\sum_{y'\in\mathbb{L}\setminus y} a_y^{y'}\log b_y^{y'}\varpi(y)\right)_+ \leq (\log\theta[0,1])_+\ell,\end{aligned}$$

where the last inequality follows on observing that $a_y^{y'}\varpi(y) \leq 1$ for all y. Thus we have that

$$\alpha \leq \exp\left\{\theta[0,1] + \frac{\ell}{e} + (\log\theta[0,1])_+\ell\right\} \equiv c_1.$$

Let $\vartheta = \alpha\tilde{\vartheta}^*$. Then $\int_{\mathbb{L}} \hat{l}(\vartheta(y,\cdot))\varpi(dy) \leq G_1(\varpi, x) + \epsilon_0$ and

$$\sup_{y\in\mathbb{L}}\left|\varpi(y)\frac{d\vartheta(y,\cdot)}{d\theta}\right| \leq c_1.$$

The result follows. □

For $\delta \in (0,1)$, let $\mathcal{P}_\delta(\mathbb{L})$ be the collection of all probability measures μ on $\mathbb{L}$ with the property that $\mu(i) \geq \delta$ for all $i \in \mathbb{L}$.

Lemma 5.5.5. *Given $\kappa \in (0,1)$ and a compact set $K_1 \subset \mathbb{R}^d$, there exists a $\delta \in (0,1)$, $B_0, L_0 \in (1,\infty)$, $l_0 \in (0,\infty)$ such that for every $\sigma \in (0,1)$, $x \in K_1$ and $v \in \mathbb{R}^d$ there is a $\varpi^* \in \mathcal{P}_\delta(\mathbb{L})$, $\tilde{v} \in \mathbb{R}^d$, $r^* \in \mathbb{R}^d$, $u^* \in \mathcal{U}(\varpi^*, x, \tilde{v} - r^*)$, $\vartheta^* \in \bar{\mathcal{K}}(\varpi^*, x)$, with the following properties.*

(i) $|r^*| \le B_0(1+|v|)$*; for all* $y \in \mathbb{L}$, $|u^*(y)| \le \frac{B_0(1+|v|)}{\sigma^2}$*; and*

$$|v - \tilde{v}| \le B(K_1)\kappa, \text{ where } B(K_1) = \sup_{x \in K_1, y \in \mathbb{L}} |b(x,y)|.$$

(ii) For all $y \in \mathbb{L}$,

$$l_0 \le \frac{d\vartheta^*(y,\cdot)}{d\theta}(w) \le L_0, \ a.s.\, \theta. \tag{5.5.11}$$

(iii) For all $\tilde{x} \in \mathbb{R}^d$ *and* $y, y' \in \mathbb{L}$,

$$\hat{c}^{\vartheta^*(y,\cdot)}(\tilde{x}, y) \ge \delta\underline{\lambda}, \ \sum_{n=1}^{\ell} \hat{R}_n^{\vartheta^*(y,\cdot)}(\tilde{x}, y, y') \ge f(\delta, L_0)\alpha \tag{5.5.12}$$

where $\hat{R}_n^{\vartheta^*(y,\cdot)}$ *is defined as in* (5.2.5) *on replacing* R *with* $\hat{R}^{\vartheta^*(y,\cdot)}$*, and* $f(\delta, L_0) = \left(\frac{\delta}{L_0}\right)^l$.

(iv) The following inequality holds

$$G_\sigma(\varpi^*, x, \tilde{v}) \le \int_{\mathbb{L}} \left(\frac{1}{2}|u^*(y)|^2 + \hat{l}(\vartheta^*(y,\cdot))\right)\varpi^*(dy) + \frac{|r^*|^2}{2\sigma^2} \le L_\sigma(x,v)(1+\kappa)+\kappa.$$

Proof. Fix $\kappa \in (0,1)$ and a compact set K_1 in $\mathbb{R}^d$. Choice of δ_0, l_0, L_0 will be specified later in the proof. Given $x \in K_1$, $v \in \mathbb{R}^d$ and $\sigma \in (0,1)$ choose $\varpi \in \mathcal{P}(\mathbb{L})$ such that

$$G_\sigma(\varpi, x, v) \le L_\sigma(x,v) + \frac{\kappa}{4}. \tag{5.5.13}$$

From Lemma 5.5.3, we can find a $r^* \in \mathbb{R}^d$ and $u \in \mathcal{U}(\varpi, x, v - r^*)$, such that with B as in Lemma 5.5.3, $|r^*| \le B(1+|v|)$ and $|u(i)| \le \frac{B(1+|v|)}{\sigma^2}$ for all $i \in \mathbb{L}$ and such that

$$\frac{1}{2}\int |u(y)|^2 \varpi(dy) + \frac{|r^*|^2}{2\sigma^2} = G_{2,\sigma}(\varpi, x, v). \tag{5.5.14}$$

Fix $\gamma \le \kappa/2$ and $\delta = \gamma\underline{\rho}$. Using Lemma 5.5.4 choose $\vartheta \in \bar{\mathcal{K}}(\varpi, x)$ such that

$$\int_{\mathbb{L}} \hat{l}(\vartheta(y,\cdot))\varpi(dy) \le G_1(\varpi, x) + \frac{\kappa}{4} \text{ and } \max_{y \in \mathbb{L}} \varpi(y)\frac{d\vartheta(y,\cdot)}{d\theta} \le c_1. \tag{5.5.15}$$

Define $\zeta \in \mathcal{P}(\mathbb{L} \times \mathcal{M}_F)$ as

$$\zeta(dyd\eta) = (1-\gamma)\delta_{\vartheta(y,\cdot)}(d\eta)\varpi(dy) + \gamma\delta_\theta(d\eta)\rho_x(dy).$$

Then $\varpi^* \in \mathcal{P}(\mathbb{L})$ defined as

$$\varpi^*(dy) = \int_{\mathcal{M}_F} \zeta(dyd\eta) = (1-\gamma)\varpi(dy) + \gamma\rho_x(dy)$$

satisfies $\varpi^*(i) \geq \gamma\rho_x(i) \geq \delta$ for all $i \in \mathbb{L}$, i.e. $\varpi^* \in \mathcal{P}_\delta(\mathbb{L})$. Disintegrating ζ as

$$\zeta(dyd\eta) = q^*(y, d\eta)\varpi^*(dy),$$

we see that $q^* \in \mathcal{K}(\varpi^*, x)$ and consequently $\vartheta^* \equiv \vartheta^{q^*} \in \bar{\mathcal{K}}(\varpi^*, x)$. Also, from (5.5.15)

$$\begin{aligned}\int_{\mathbb{L}} \hat{l}(\vartheta^*(y,\cdot))\varpi^*(dy) \leq &\int_{\mathbb{L}}\int_{\mathcal{M}_F} \hat{l}(\eta)q^*(y,d\eta)\varpi^*(dy) \\ =&(1-\gamma)\int_{\mathbb{L}} \hat{l}(\vartheta(y,\cdot))\varpi(dy) \\ \leq&G_1(\varpi, x) + \frac{\kappa}{4}. \end{aligned} \tag{5.5.16}$$

Next, we claim that

$$\delta \leq \frac{d\vartheta^*(y,\cdot)}{d\theta}(w) \leq \frac{c_1+1}{\delta}, \text{ a.s. } \theta. \tag{5.5.17}$$

Indeed, for $A \in \mathcal{B}[0,1]$,

$$\begin{aligned}\vartheta^*(y, A) =& \int_{\mathcal{M}_F} \eta(A)q^*(y,d\eta) \leq \frac{1}{\delta}\int_{\mathcal{M}_F} \eta(A)q^*(y,d\eta)\varpi^*(y) \\ =&\frac{1}{\delta}\left((1-\gamma)\vartheta(y,A)\varpi(y) + \gamma\theta(A)\rho_x(y)\right) \leq \frac{1}{\delta}\left((1-\gamma)c_1\theta(A) + \gamma\theta(A)\right) \\ \leq&\frac{c_1+1}{\delta}\theta(A).\end{aligned}$$

Similarly

$$\begin{aligned}\vartheta^*(y, A) =& \int_{\mathcal{M}_F} \eta(A)q^*(y,d\eta) \geq \int_{\mathcal{M}_F} \eta(A)q^*(y,d\eta)\varpi^*(y) \\ =&\left((1-\gamma)\vartheta(y,A)\varpi(y) + \gamma\theta(A)\rho_x(y)\right) \\ \geq&\delta\theta(A).\end{aligned}$$

This proves (5.5.17) and thus part (ii) follows with $L_0 = \frac{c_1+1}{\delta}$ and $l_0 = \delta$.

For $(\tilde{x}, y) \in \mathbb{G}$

$$\hat{c}^{\vartheta^*(y,\cdot)}(\tilde{x}, y) = \vartheta^*(y, E_{\tilde{x},y}) = \frac{1}{\varpi^*(y)}\vartheta^*(y, E_{\tilde{x},y})\varpi^*(y) \geq \frac{\gamma}{\varpi^*(y)}\theta(E_{\tilde{x},y})\rho_x(y) \geq \delta\underline{\lambda}.$$

Similarly one sees that

$$\sum_{n=1}^{\ell} \hat{R}_n^{\vartheta^*(y,\cdot)}(\tilde{x}, y, y') \geq f(\delta, L_0)\alpha, \quad \text{for all } \tilde{x} \in \mathbb{R}^d.$$

Thus (iii) is satisfied.

Next note that

$$\int_{\mathbb{L}} a(x,y)u(y)\varpi(dy) = \int_{\mathbb{L}} a(x,y)u^*(y)\varpi^*(dy),$$

where $u^*(y) = u(y)\frac{\varpi(y)}{\varpi^*(y)}$. Also

$$\int_{\mathbb{L}} a(x,y)u^*(y)\varpi^*(dy) + \int_{\mathbb{L}} b(x,y)\varpi^*(dy) = v - r^* + \int_{\mathbb{L}} b(x,y)(\varpi^*(dy) - \varpi(dy)) = \tilde{v} - r^*,$$

where $\tilde{v} = v + \int_{\mathbb{L}} b(x,y)(\varpi^*(dy) - \varpi(dy))$. Thus $u^* \in \mathcal{U}(\varpi^*, x, \tilde{v} - r^*)$. Also

$$|\tilde{v} - v| \leq \kappa \sup_{x \in K_1, y \in \mathbb{L}} |b(x,y)| = \kappa B(K_1) \text{ and } |u^*(y)| \leq \frac{B(1+|v|)}{\delta\sigma^2}.$$

This proves (i) with $B_0 = \frac{B}{\delta}$. Next,

$$\frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2\varpi^*(dy) = \frac{1}{2}\int_{\mathbb{L}} |u(y)|^2\frac{\varpi(y)}{\varpi^*(y)}\varpi(dy) \leq \frac{1}{1-\gamma}\frac{1}{2}\int_{\mathbb{L}} |u(y)|^2\varpi(dy).$$

Thus from (5.5.13) and (5.5.14)

$$\begin{aligned}\frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2\varpi^*(dy) + \frac{|r^*|^2}{2\sigma^2} \leq& \frac{1}{2}\int_{\mathbb{L}} |u(y)|^2\varpi(dy) + \frac{|r^*|^2}{2\sigma^2} + \frac{\gamma}{1-\gamma}(L_\sigma(x,v) + \frac{\kappa}{4}) \\ \leq& G_{2,\sigma}(\varpi, x, v) + \frac{\kappa}{4} + \kappa L_\sigma(x,v),\end{aligned}$$

where the last inequality follows from our choice of γ.

Combining the estimate in the above display with (5.5.13) and (5.5.16)

$$
\begin{aligned}
\frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2 \varpi^*(dy) + & \int_{\mathbb{L}} \hat{l}(\ \vartheta^*(y,\cdot))\varpi^*(dy) + \frac{|r^*|^2}{2\sigma^2} \\
& \le G_1(\varpi, x) + \frac{\kappa}{4} + G_{2,\sigma}(\varpi, x, v) + \frac{\kappa}{4} + \kappa L_\sigma(x,v) \\
& = G_\sigma(\varpi, x, v) + \frac{\kappa}{2} + \kappa L_\sigma(x,v) \\
& \le L_\sigma(x,v) + \kappa + \kappa L_\sigma(x,v).
\end{aligned}
$$

This proves (iv). □

Proof of the following lemma is given in the Appendix.

Lemma 5.5.6. *For every $m \in (0,\infty)$ and $\delta \in (0,1)$ there exists a $\beta(m,\delta) \in (0,\infty)$ such that whenever $\vartheta^* \in \bar{\mathcal{K}}$ satisfies* (5.5.12) *and*

$$\max_{y\in\mathbb{L}} \frac{d\vartheta^*(y,\cdot)}{d\theta}(w) \le m \quad a.s.\ \theta, \tag{5.5.18}$$

we have

$$|\varpi_x - \varpi_{x'}| \le \beta(m,\delta)|x - x'|, \ \textit{for all } x, x' \in \mathbb{R}^d$$

where ϖ_x is the unique invariant measure of the Markov process with generator $\Pi_x^{\vartheta^*}$.

Lemma 5.5.7. *For every $\sigma \in (0,\infty)$, L_σ is continuous on $\mathbb{R}^d \times \mathbb{R}^d$.*

Proof. Fix $\sigma \in (0,\infty)$. We will show that for every compact $K = K_1 \times K_2 \subset \mathbb{R}^d \times \mathbb{R}^d$ and $\kappa_0 \in (0,1)$ there exists a $\epsilon_0 \in (0,1)$ such that

$$L_\sigma(x',v') \le L_\sigma(x,v) + \kappa_0, \quad \text{whenever } (x,v),(x',v') \in K \text{ and } |x-x'| + |v-v'| \le \epsilon_0.$$

Fix such a κ_0 and K. Let $M_1 = \sup_{(x,v)\in K} L_\sigma(x,v)$. Note that M_1 is finite from Lemma 5.5.2. We will specify a choice of ϵ_0 later in the proof. Fix $(x,v) \in K$. We now apply Lemma 5.5.5 with K_1 and $\kappa = \min\left\{\frac{\kappa_0}{2(1+M_1)}, \frac{\kappa_0\sigma^2}{32B_1B(K_1)}\right\}$. Then there are $\delta \in$

$(0,1)$, $L_0, B_0 \in (0,\infty)$, $\varpi^* \in \mathcal{P}_\delta(\mathbb{L})$, $\vartheta^* \in \bar{\mathcal{K}}(\varpi^*, x)$, $r^* \in \mathbb{R}^d$ and $u^* \in \mathcal{U}(\varpi^*, x, \tilde{v} - r^*)$ such that parts (i) - (iv) of Lemma 5.5.5 are satisfied. Since $\varpi^* \in \mathcal{P}_\delta(\mathbb{L})$, we have that

$$\max_{y\in\mathbb{L}} \hat{l}(\vartheta^*(y,\cdot)) \leq \frac{M_1+1}{\delta} \equiv M_2.$$

Also, using Lemma 5.5.6

$$\max_{i\in\mathbb{L}} |\varpi^*(i) - \varpi^*_{x'}(i)| \leq \beta(L_0,\delta)|x-x'|$$

for all $x' \in K_1$, where $\varpi^*_{x'}$ is the unique invariant distribution for $\Pi^{\vartheta^*}_{x'}$. Let $B_1 = \sup_{v\in K_2} B_0(1+|v|)$ Choose ϵ_0 small enough so that whenever $|x-x'|+|v-v'| \leq \epsilon_0$

$$\max_{i\in\mathbb{L}} |\varpi^*(i) - \varpi^*_{x'}(i)| \leq \frac{\kappa_0}{4M_2\ell} \wedge \frac{\kappa_0\sigma^4}{4B_1^2\ell} \tag{5.5.19}$$

and

$$\left|(v-v') + \int (a(x',y)u^*(y) + b(x',y))\varpi^*_{x'}(dy) - \int (a(x,y)u^*(y)+b(x,y))\varpi^*(dy)\right| \leq \frac{\kappa_0\sigma^2}{32B_1}.$$

Since $\vartheta^* \in \bar{\mathcal{K}}(\varpi^*_{x'}, x')$, we have

$$\begin{aligned}
G_1(\varpi^*_{x'}, x') \leq& \int_{\mathbb{L}} \hat{l}(\vartheta^*(y,\cdot))(\varpi^*_{x'} - \varpi^*)(dy) + \int_{\mathbb{L}} \hat{l}(\vartheta^*(y,\cdot))\varpi^*(dy) \\
\leq& \frac{\kappa_0}{4M_2\ell} M_2\ell + \int_{\mathbb{L}} \hat{l}(\vartheta^*(y,\cdot))\varpi^*(dy) \\
\leq& \frac{\kappa_0}{4} + \int_{\mathbb{L}} \hat{l}(\vartheta^*(y,\cdot))\varpi^*(dy)
\end{aligned} \tag{5.5.20}$$

Next, define

$$r_0 = (\tilde{v} - v') + \int (a(x',y)u^*(y) + b(x',y))\varpi^*_{x'}(dy) - \int (a(x,y)u^*(y) + b(x,y))\varpi^*(dy),$$

note that $|r_0| \leq \frac{\kappa_0\sigma^2}{16B_1}$. Let $\tilde{r} = r^* - r_0$, then $u^* \in \mathcal{U}(\varpi^*_{x'}, x', v' - \tilde{r})$. Also,

$$\begin{aligned}
G_{2,\sigma}(\varpi^*_{x'}, x', v') \leq& \frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2 \varpi^*_{x'}(dy) + \frac{|\tilde{r}|^2}{2\sigma^2} \\
\leq& \frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2 \varpi^*_{x'}(dy) + \frac{|r^*|^2}{2\sigma^2} + \frac{|r_0|^2}{2\sigma^2} + \frac{|r^*||r_0|}{\sigma^2} \\
\leq& \frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2 \varpi^*_{x}(dy) + \frac{\kappa_0}{8} + \frac{|r^*|^2}{2\sigma^2} + \frac{\kappa_0}{16} + \frac{\kappa_0}{16}.
\end{aligned} \tag{5.5.21}$$

Combining estimates in (5.5.19), (5.5.20), (5.5.21), Lemma 5.5.5(iv)

$$\begin{aligned}
L_\sigma(x', v') \leq & G_\sigma(\varpi^*_{x'}, x', v') \\
\leq & \frac{\kappa_0}{2} + \frac{1}{2}\int_{\mathbb{L}} |u^*(y)|^2 \varpi^*(dy) + \int_{\mathbb{L}} \hat{l}(\vartheta^*(y, \cdot))\varpi^*(dy) + \frac{|r^*|^2}{2\sigma^2} \\
\leq & L_\sigma(x, v) + \frac{\kappa_0}{2} + \kappa(1 + L_\sigma(x, v)) \\
\leq & L_\sigma(x, v) + \kappa_0.
\end{aligned}$$

The result follows. □

5.5.2 Auxiliary Lemmas.

In this section we collect two other lemmas that will be used in the proof of the lower bound. Let $\mathrm{PL}_{x_0}([0, T] : \mathbb{R}^d)$ be the space of piecewise linear maps from $[0, T]$ to $\mathbb{R}^d$ starting from x_0, i.e. $\varphi \in \mathbb{U}$ belongs to $\mathrm{PL}_{x_0}([0, T] : \mathbb{R}^d)$ if it is absolutely continuous, $\varphi(0) = x_0$ and there exists a partition $0 = t_0 < t_1 \cdots < t_k = T$ such that $\dot{\varphi}(t) = \dot{\varphi}(t_i+)$ for all $t \in (t_i, t_{i+1})$, $i = 0, \cdots k-1$. For $\xi \in \mathbb{U}$, let $||\xi||_T = \sup_{0\leq t\leq T} |\xi(t)|$.

Lemma 5.5.8. *Let $\varphi \in \mathbb{U}$, $\varphi(0) = x_0$ be such that φ is absolutely continuous and $\int_0^T |\dot{\varphi}(s)|^2 ds < \infty$. Fix $\sigma > 0$. Then there exists a sequence $\{\varphi_n\} \subset PL_{x_0}([0, T] : \mathbb{R}^d)$ such that $\varphi_n \to \varphi$ uniformly and, as $n \to \infty$*

$$\int_0^T L_\sigma(\varphi_n(t), \dot{\varphi}_n(t))dt \to \int_0^T L_\sigma(\varphi(t), \dot{\varphi}(t))dt.$$

Proof. Let $\psi = \dot{\varphi}$ and define, for $\delta > 0$, $\psi^\delta(s) = \delta^{-1}\int_{(s-\delta)_+}^s \psi(u)du$, and $\varphi^\delta(s) = \int_0^s \psi^\delta(u)du$, $s \in [0, T]$. Then $s \to \psi^\delta(s)$ is continuous and

$$\int_0^T |\psi^\delta(s)|^2 ds \to \int_0^T |\psi(s)|^2 ds, \text{ as } \delta \to 0. \tag{5.5.22}$$

Also

$$(\varphi^\delta(s), \psi^\delta(s)) \to (\varphi(s), \psi(s)) \text{ for every } s \in [0, T] \tag{5.5.23}$$

and the convergence of φ^δ to φ is uniform. Thus, from Lemma 5.5.7,

$$L_\sigma(\varphi^\delta(s), \psi^\delta(s)) \to L_\sigma(\varphi(s), \psi(s)) \text{ as } \delta \to 0. \tag{5.5.24}$$

Combining this fact with Lemma 5.5.2 and (5.5.22) we have that

$$\int_0^T L_\sigma(\varphi^\delta(t), \psi^\delta(t))dt \to \int_0^T L_\sigma(\varphi(t), \psi(t))dt.$$

Thus, in order to prove the lemma, we can assume without loss of generality that $\dot{\varphi}$ is continuous. In particular $\sup_{s\in[0,T]}\{|\varphi(s)| + |\dot{\varphi}(s)|\} \equiv \tilde{M}_0 < \infty$. Let $K = \{(x, v) : |x| + |v| \leq \tilde{M}_0\}$. Fix $\kappa > 0$. From Lemma 5.5.7, we can find $\delta_0 \in (0, \kappa)$ such that whenever $(x, v), (x', v') \in K$ and $|x - x'| + |v - v'| \leq \delta_0$, we have

$$|L_\sigma(x, v) - L_\sigma(x', v')| \leq \epsilon_0 = \frac{\kappa}{T}.$$

Also, from continuity of $\varphi, \dot{\varphi}$, we can find $\gamma_0 \in (0, \frac{\delta_0}{2\tilde{M}_0})$ such that whenever $s, t \in [0, T]$, $|s - t| \leq \gamma_0$, we have

$$|\varphi(s) - \varphi(t)| \leq \frac{\delta_0}{4} \text{ and } |\dot{\varphi}(s) - \dot{\varphi}(t)| \leq \frac{\delta_0}{4}.$$

Now consider the partition $0 = t_0 < t_1 \cdots < t_k = T$ where, with $k = \lfloor T/\gamma_0 \rfloor + 1$, $t_i = i\gamma_0$, $i = 0, 1, \cdots k-1$ and $t_k = T$. Define φ^κ such that $\varphi^\kappa(t_i) = \varphi(t_i), i = 0, 1, \cdots k-1$ and is extended to $[0, T]$ by the linear interpolation of the points $\varphi^\kappa(t_i), i = 0, 1, \cdots k-1$. In particular,

$$\dot{\varphi}^\kappa(t) = \dot{\varphi}(\tilde{t}_i), \text{ for some } \tilde{t}_i \in (t_i, t_{i+1}) \text{ whenever } t \in (t_i, t_{i+1}), \ i = 0, 1, \cdots k-1.$$

Then $\varphi^\kappa \in \mathrm{PL}_{x_0}([0, T] : \mathbb{R}^d)$ and

$$\begin{aligned}
&\left|\int_0^T L_\sigma(\varphi^\kappa(t), \dot{\varphi}^\kappa(t))dt - \int_0^T L_\sigma(\varphi(t), \dot{\varphi}(t))dt\right| \\
\leq &\sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \left|L_\sigma(\varphi^\kappa(t), \dot{\varphi}^\kappa(\tilde{t}_i)) - L_\sigma(\varphi(t), \dot{\varphi}(t))\right| dt \\
\leq &\epsilon_0 T = \kappa
\end{aligned}$$

where the last inequality follows on observing that for all $s \in [t_i, t_{i+1})$, $i = 0, 1, \cdots k-1$,

$$
\begin{aligned}
|\varphi^\kappa(s) - \varphi(s)| + |\dot\varphi^\kappa(s) - \dot\varphi(s)| \leq & |\varphi(t_i) - \varphi(s)| + |s - t_i||\dot\varphi(\tilde t_i)| + |\dot\varphi(s) - \dot\varphi(\tilde t_i)| \\
\leq & \frac{\delta_0}{2} + \gamma_0 \tilde M_0 \leq \frac{\delta_0}{2} + \frac{\delta_0}{2\tilde M_0}\tilde M_0 = \delta_0.
\end{aligned}
$$

Also, since $\delta_0 \leq \kappa$, we have $\sup_{s\in[0,T]} |\varphi^\kappa(s) - \varphi(s)| \leq \kappa$. Thus for an arbitrary $\kappa > 0$, we have found a $\varphi^\kappa \in \mathrm{PL}_{x_0}([0,T] : \mathbb{R}^d)$ such that

$$
\sup_{s\in[0,T]} |\varphi^\kappa(s) - \varphi(s)| \leq \kappa, \; \left| \int_0^T L_\sigma(\varphi^\kappa(t), \dot\varphi^\kappa(t))dt - \int_0^T L_\sigma(\varphi(t), \dot\varphi(t))dt \right| \leq \kappa.
$$

The result follows. $\square$

In fact in the course of the proof of the above lemma we have established the following result.

Lemma 5.5.9. *Let $\varphi \in \mathbb{U}$, $\varphi(0) = x_0$ be such that φ is absolutely continuous and $\int_0^T |\dot\varphi(s)|^2 ds < \infty$. Fix $\sigma \in (0,1)$. Then, for every $\kappa > 0$ there exists a $\varphi^\kappa \in PL_{x_0}([0,T] : \mathbb{R}^d)$ such that $||\varphi^\kappa - \varphi||_T \leq \kappa$, and if $0 = t_0 < t_1 \cdots < t_k = T$ is the associated partition then for all $s_i \in [t_i, t_{i+1}]$, $i = 0, 1, \cdots k-1$,*

$$
\left| \sum_{i=0}^{k-1} L_\sigma(\varphi^\kappa(s_i), \dot\varphi^\kappa(s_i))(t_{i+1} - t_i) - \int_0^T L_\sigma(\varphi(t), \dot\varphi(t))dt \right| \leq \kappa.
$$

5.5.3 Proof of the lower bound.

We now prove the inequality (5.5.1). We can assume without loss of generality that F is a Lipschitz function (See Corollary 1.2.5 in [33]). Let $M_F \in (0, \infty)$ be such that

$$
|F(\xi) - F(\tilde\xi)| \leq M_F \sup_{0\leq t\leq T} |\xi(t) - \tilde\xi(t)|, \text{ for all } \xi, \tilde\xi \in \mathbb{U}. \tag{5.5.25}
$$

From Assumption 5.2.1(2) we have that for some $d_{lip} \in (0, \infty)$

$$\max_{y \in \mathbb{L}} \{|a(x, y) - a(x', y)| + |b(x, y) - b(x', y)|\} \leq d_{lip}|x - x'|, \quad \text{for all } x, x' \in \mathbb{R}^d. \tag{5.5.26}$$

Fix $\epsilon_0 \in (0, 1)$. Choose $\xi_1 \in \mathbb{U}$ such that

$$F(\xi_1) + \mathbb{I}(\xi_1) \leq \inf_{\xi \in \mathbb{U}} \{F(\xi) + \mathbb{I}(\xi)\} + \frac{\epsilon_0}{2}.$$

From Proposition 5.5.1, $\mathbb{I}(\xi_1) \geq \tilde{\mathbb{I}}(\xi_1)$. Therefore, since $L_\sigma(x, v) \leq L(x, v)$,

$$F(\xi_1) + \int_0^T L_\sigma(\xi_1(s), \dot{\xi}_1(s))ds \leq \inf_{\xi \in \mathbb{U}} \{F(\xi) + \mathbb{I}(\xi)\} + \frac{\epsilon_0}{2}.$$

Now fix a $\sigma \in (0, 1)$. Using Lemma 5.5.9, we can now find a $\xi^* \in \mathrm{PL}_{x_0}([0, T] : \mathbb{R}^d)$ such that $||\xi_1 - \xi^*||_T \leq \epsilon_0$ and

$$F(\xi^*) + \sum_{i=0}^{k-1} L_\sigma(\xi^*(s_i), \dot{\xi}^*(t_i))(t_{i+1} - t_i) \leq \inf_{\xi \in \mathbb{U}} \{F(\xi) + \mathbb{I}(\xi)\} + \epsilon_0, \tag{5.5.27}$$

for all $s_i \in [t_i, t_{i+1}]$, where $0 = t_0 < t_1 \cdots < t_k = T$ is the partition over which ξ^* is piecewise linear. From the upper bound (5.4.1) proved in Section 5.4 we have that

$$\sum_{i=0}^{k-1} L_\sigma(\xi^*(s_i), \dot{\xi}^*(t_i))(t_{i+1} - t_i) \leq 2||F||_\infty + 1 \equiv M_0, \tag{5.5.28}$$

Let $K_1 = \{x \in \mathbb{R}^d : |x| \leq \bar{M}\}$, where $\bar{M} = ||\xi_1||_T + 1$. We apply Lemma 5.5.5 with the compact set K_1 and $\kappa = \epsilon_0$. Let δ, L_0, l_0 be as in Lemma 5.5.5. Define $M < \infty$ as

$$M = \sup_{x : |x| \leq \bar{M}} \sup_{y \in \mathbb{L}} [|b(x, y)| + |a(x, y)|]. \tag{5.5.29}$$

Choose

$$d_0 = \frac{\epsilon_0}{2} \min\left\{1, \frac{1}{M_F}, \frac{\delta}{2(2M_0 + 1)\beta(L_0, \delta)}\right\}, \tag{5.5.30}$$

where β is as in Lemma 5.5.6. Let

$$M_1 = (d_{lip} + M\beta(L_0, \delta))\left(T + \frac{\sqrt{2}}{\delta}(2M_0 + T)\right) + 1.$$

From (5.5.27) we have that

$$F(\xi^*) + \sum_{i=0}^{k-1} L_\sigma(\xi^*(t_i), \dot{\xi}^*(t_i))(t_{i+1} - t_i) \le \inf_{\xi \in \mathbb{U}} \{F(\xi) + \mathbb{I}(\xi)\} + \epsilon_0, \tag{5.5.31}$$

where, by refining the partition if necessary, we can assume without loss of generality that $\max_i |t_{i+1} - t_i| = \Lambda$ is such that

$$\sup_{s,t \in [0,T], |t-s| \le \Lambda} |\xi^*(s) - \xi^*(t)| = \bar{\Lambda} \le \frac{d_0}{M_1 \exp(M_1)}, \quad \text{and } k \le \frac{T}{\Lambda} + 1. \tag{5.5.32}$$

For $j = 0, 1, \cdots (k-1)$, let

$$m_j = 2L_\sigma(\xi^*(t_j), \dot{\xi}^*(t_j)) + 1, \; \Lambda_j = (t_{j+1} - t_j) \text{ and } D_j = (1 + \sqrt{2} m_j/\delta)(d_{lip} + M\beta(L_0, \delta)). \tag{5.5.33}$$

Then

$$\sum_{j=0}^{k-1} D_j \Lambda_j \le M_1. \tag{5.5.34}$$

Then from Lemma 5.5.5, for each $j \in \{0, 1, \cdots k-1\}$ there is a

$$\tilde{v}_j \in \mathbb{R}^d, \; \varpi_j^* \in \mathcal{P}_\delta(\mathbb{L}), \; \vartheta_j \in \bar{\mathcal{K}}(\varpi_j^*, \xi^*(t_j)), \; r_j \in \mathbb{R}^d, \; u_j \in \mathcal{U}(\varpi_j^*, \xi^*(t_j), \tilde{v}_j - r_j)$$

such that for all $y \in \mathbb{L}$

$$l_0 \le \frac{d\vartheta_j(y, \cdot)}{d\theta}(w) \le L_0, \text{ a.s. } \theta \tag{5.5.35}$$

$|\tilde{v}_j - \dot{\xi}^*(t_j)| \le B(K_1)\epsilon_0$, for all $j = 0, 1, \cdots k$, and the inequalities in Lemma 5.5.5 (ii) hold with ϑ^* replaced by ϑ_j, and

$$\int_{\mathbb{L}} \left(\hat{l}(\vartheta_j(y, \cdot)) + \frac{1}{2}|u_j(y)|^2 \right) \varpi_j^*(dy) + \frac{1}{2\sigma^2}|r_j|^2 \le (1 + \epsilon_0) L_\sigma(\xi^*(t_j), \dot{\xi}^*(t_j)) + \epsilon_0. \tag{5.5.36}$$

Also note that

$$\max \left(\frac{1}{2} \int_{\mathbb{L}} |u_j(y)|^2 \varpi_j^*(dy), \frac{1}{\sqrt{2}} \int_{\mathbb{L}} |u_j(y)| \varpi_j^*(dy) \right) \le m_j. \tag{5.5.37}$$

Denote the unique invariant measure of the Markov process with generator $\Pi_x^{\vartheta_j}$ as ρ_x^j. Then, from Lemma 5.5.6,

$$\max_{j=0,1,\cdots k-1} \max_{y \in \mathbb{L}} |\rho_x^j(y) - \rho_{x'}^j(y)| \le \beta(L_0, \delta)|x - x'|, \text{ for all } x, x' \in \mathbb{R}^d. \tag{5.5.38}$$

Let $\varphi_j^*(y,r) = \frac{d\vartheta_j(y,\cdot)}{d\theta}(r),\ r \in [0,1]$ and let B be a d-dimensional standard Brownian motion that is independent of the driving noises $(W, \bar{N})$. Define stochastic process $(\tilde{X}^\epsilon, \tilde{Y}^\epsilon, \tilde{U}^\epsilon)$, iteratively as follows. Let $(\tilde{X}^\epsilon(0), \tilde{Y}^\epsilon(0)) = (x_0, y_0)$ and $\tilde{U}^\epsilon(0) = 0$. Having defined $(\tilde{X}^\epsilon(t), \tilde{Y}^\epsilon(t), \tilde{U}^\epsilon(t))$ for $t \in [0, t_j]$, define for $t \in (t_j, t_{j+1}]$,

$$\begin{aligned}
\tilde{X}^\epsilon(t) =& \tilde{X}^\epsilon(t_j) + \int_{t_j}^t b(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))ds + \sqrt{\epsilon}\int_{t_j}^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))dW(s) \\
&+ \int_{t_j}^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))\psi_j^\epsilon(s)ds \\
\tilde{Y}^\epsilon(t) =& \tilde{Y}^\epsilon(t_j) + \int_{t_j}^t \int_{r\in[0,1]} k(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s-), r) N^{\frac{1}{\epsilon}\varphi_j^\epsilon}(dr \times ds), \\
\tilde{U}^\epsilon(t) =& \tilde{U}^\epsilon(t_j) + \sqrt{\epsilon}\sigma \int_{t_j}^t dB(s) + r_j(t - t_j),
\end{aligned} \tag{5.5.39}$$

where

$$\psi_j^\epsilon(t) = u_j(\tilde{Y}^\epsilon(t)),\ \varphi_j^\epsilon(t,r) = \varphi_j^*(\tilde{Y}^\epsilon(t-), r), \tag{5.5.40}$$

Finally, let $\tilde{Z}^\epsilon = \tilde{X}^\epsilon + \tilde{U}^\epsilon$. Following the proof of Theorem II.8 in [73], we see that as $\epsilon \to 0$, $\tilde{X}^\epsilon$ converges in probability (with the uniform metric on $\mathbb{U}$) to $\hat{x} \in \mathbb{U}$ given as the unique solution of the equation

$$\begin{aligned}
d\hat{x}(t) =& \left(\int_{\mathbb{L}} \left(b(\hat{x}(t), y) + a(\hat{x}(t), y)u_j(y) \right) \rho^j_{\hat{x}(t)}(dy) \right) dt, \\
t \in& [t_j, t_{j+1}],\ j = 0, 1, \cdots, k-1,\ \hat{x}(0) = x_0.
\end{aligned} \tag{5.5.41}$$

Also, $\tilde{U}^\epsilon$ converges in probability (with uniform metric on $\mathbb{U}$) to ζ, defined as

$$\zeta(0) = 0;\ \zeta(t) = \zeta(t_j) + r_j(t - t_j),\ t \in (t_j, t_{j+1}],\ j = 0, 1, \cdots, k-1. \tag{5.5.42}$$

Let $z = \hat{x} + \zeta$. Then $\tilde{Z}^\epsilon \to z$ in probability as $\epsilon \to 0$.

Next note that

$$\begin{aligned}
|\zeta|_T^2 \leq& \sup_{0\leq t\leq T} \left(\int_0^t \left| \dot{\zeta}(s) \right| ds \right)^2 \leq T \int_0^T |\dot{\zeta}(s)|^2 ds = 2T\sigma^2 \sum_{j=0}^{k-1} \frac{|r_j|^2}{2\sigma^2}(t_{j+1} - t_j) \\
\leq& 2T\sigma^2 \left(2\sum_{j=0}^{k-1} L_\sigma(\xi^*(t_j), \dot{\xi}^*_{t_j})(t_{j+1} - t_j) + 1 \right) \leq 2T\sigma^2(2M_0 + 1) \equiv (s(\sigma))^2,
\end{aligned} \tag{5.5.43}$$

where the next to last inequality follows from (5.5.36) and the last inequality follows from (5.5.28).

Next define $\varrho \in \mathbb{U}$ as

$$\varrho(0) = 0;\ \varrho(t) = \varrho(t_j) - [\tilde{v}_j - \dot{\xi}^*(t_j)](t - t_j),\ t \in (t_j, t_{j+1}],\ j = 0, 1, \cdots, k-1$$

Note that

$$\int_s^t |\dot{\varrho}(u)| du \leq B(K_1)\epsilon_0(t - s),\ \text{for all } 0 \leq s \leq t \leq T. \tag{5.5.44}$$

We now estimate $\sup_{0 \leq t \leq T} |\hat{x}(t) - \xi^*(t)|$. Note that for $t \in [0, t_1]$

$$\begin{aligned}
\hat{x}(t) - \xi^*(t) = &\int_0^t \left[\int_{\mathbb{L}} b(\hat{x}(s), y) \rho^0_{\hat{x}(s)}(dy) - \int_{\mathbb{L}} b(\xi^*_0, y) \varpi^*_0(dy) \right] ds \\
&+ \int_0^t \left[\int_{\mathbb{L}} a(\hat{x}(s), y) u_0(y) \rho^0_{\hat{x}(s)}(dy) - \int_{\mathbb{L}} a(\xi^*_0, y) u_0(y) \varpi^*_0(dy) \right] ds \\
&- \zeta(t) + \varrho(t).
\end{aligned}$$

For the second term, we have for $t \in [0, t_1]$,

$$\begin{aligned}
&\left| \int_{\mathbb{L}} a(\hat{x}(t), y) u_0(y) \rho^0_{\hat{x}(t)}(dy) - \int_{\mathbb{L}} a(\xi^*(0), y) u_0(y) \varpi^*_0(dy) \right| \\
&\leq \left| \int_{\mathbb{L}} \left(a(\hat{x}(t), y) - a(\xi^*(0), y) \right) u_0(y) \rho^0_{\hat{x}(t)}(dy) \right| \\
&\qquad + \left| \int_{\mathbb{L}} a(\xi^*(0), y) u_0(y) \left(\rho^0_{\hat{x}(t)}(dy) - \varpi^*_0(dy) \right) \right| \\
&\leq \frac{\sqrt{2} m_0}{\delta} \left(d_{lip} + M\beta(L_0, \delta) \right) |\hat{x}(t) - \xi^*(0)|
\end{aligned}$$

where in the last inequality, we have used (5.5.37), (5.5.29), (5.5.35), (5.5.38) and the facts that $\varpi^*_0(y) \geq \delta$ and $\varpi^*_0(dy) = \rho^0_{\xi^*(0)}(dy)$. For the first term we have similarly,

$$\left| \int_{\mathbb{L}} b(\hat{x}(t), y) \rho^0_{\hat{x}(t)}(dy) - \int_{\mathbb{L}} b(\xi^*(0), y) \varpi^*_0(dy) \right| \leq \left(d_{lip} + M\beta(L_0, \delta) \right) |\hat{x}(t) - \xi^*(0)|.$$

Combining these estimates, we have on recalling the definitions of D_j and Λ_j from

(5.5.33) and using (5.5.32), for $t \in [0, t_1]$,

$$
\begin{aligned}
&|\hat{x}(t) - \xi^*(t)| \\
&\le D_0 \int_0^t |\hat{x}(s) - \xi^*(0)| ds + \int_0^t |\dot{\zeta}(s)| ds + \int_0^t |\dot{\varrho}(s)| ds \\
&\le D_0 \int_0^t |\hat{x}(s) - \xi^*(s)| ds + D_0 \int_0^t |\xi^*(s) - \xi^*(0)| ds + \int_0^t |\dot{\zeta}(s)| ds + \int_0^t |\dot{\varrho}(s)| ds \\
&\le D_0 \int_0^t |\hat{x}(s) - \xi^*(s)| ds + D_0 \Lambda_0 \bar{\Lambda} + \int_0^t |\dot{\zeta}(s)| ds + \int_0^t |\dot{\varrho}(s)| ds.
\end{aligned}
$$

Then by Gronwall's inequality, we have

$$
\sup_{t_0 \le t \le t_1} |\hat{x}(t) - \xi^*(t)| \le \big(D_0 \Lambda_0 \bar{\Lambda} + H_0\big) \exp\{D_0 \Lambda_0\},
$$

where for $j = 0, \cdots (k-1)$, $H_j = \int_{[t_j, t_{j+1}]} |\dot{\zeta}(s)| ds + \int_{[t_j, t_{j+1}]} |\dot{\varrho}(s)| ds$. Similarly, we have for $t \in [t_1, t_2]$,

$$
\begin{aligned}
&|\hat{x}(t) - \xi^*(t)| \\
&\le D_1 \int_{t_1}^t |\hat{x}(s) - \xi^*(t_1)| ds + |\hat{x}(t_1) - \xi^*(t_1)| + \int_{t_1}^t |\dot{\zeta}(s)| ds + \int_{t_1}^t |\dot{\varrho}(s)| ds \\
&\le D_1 \int_{t_1}^t |\hat{x}(s) - \xi^*(s)| ds + D_1 \int_{t_1}^t |\xi^*(s) - \xi^*(t_1)| ds + H_1 + (D_0 \Lambda_0 \bar{\Lambda} + H_0) \exp\{D_0 \Lambda_0\} \\
&\le D_1 \int_{t_1}^t |\hat{x}(s) - \xi^*(s)| ds + (D_1 \Lambda_1 \bar{\Lambda} + H_1) + (D_0 \Lambda_0 \bar{\Lambda} + H_0) \exp\{D_0 \Lambda_0\} \\
&\le D_1 \int_{t_1}^t |\hat{x}(s) - \xi^*(s)| ds + ((D_0 \Lambda_0 + D_1 \Lambda_1) \bar{\Lambda} + H_0 + H_1) \exp\{D_0 \Lambda_0\}.
\end{aligned}
$$

Thus by Gronwall's inequality, we have

$$
\sup_{t_1 \le t \le t_2} |\hat{x}(t) - \xi^*(t)| \le ((D_0 \Lambda_0 + D_1 \Lambda_1) \bar{\Lambda} + H_0 + H_1) \exp\{D_0 \Lambda_0 + D_1 \Lambda_1\}.
$$

Using similar estimates recursively, we have from (5.5.34), (5.5.43), (5.5.44) and (5.5.32)

$$
\begin{aligned}
\sup_{0 \le t \le T} |\hat{x}(t) - \xi^*(t)| &\le \left(\bar{\Lambda} \sum_{j=0}^{k-1} D_j \Lambda_j + \sum_{j=0}^{k-1} H_j\right) \exp\left(\sum_{j=0}^{k-1} D_j \Lambda_j\right) \\
&\le \bar{\Lambda} M_1 \exp(M_1) + s(\sigma) \exp(M_1) + \epsilon_0 A_1 \\
&\le d_0 + s(\sigma) \exp(M_1) \le \frac{\epsilon_0}{2} + s(\sigma) \exp(M_1) + \epsilon_0 A_1, \qquad (5.5.45)
\end{aligned}
$$

where $A_1 = B(K_1)T\exp(M_1)$. Combining this with (5.5.32), we have

$$|\hat{x}(t) - \xi^*(t_i)| \leq 2d_0 + s(\sigma)\exp(M_1) + \epsilon_0 A_1, \text{ for all } t \in [t_i, t_{i+1}], i = 0, 1, \cdots, k-1. \tag{5.5.46}$$

By standard arguments it follows that, as $\epsilon \to 0$,

$$\begin{aligned} \frac{1}{2}\int_0^T |\psi^\epsilon(s)|^2 ds &\to \sum_{i=0}^{k-1} \frac{1}{2}\int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} |u_i(y)|^2 \rho^i_{\hat{x}(s)}(dy)ds, \\ \int_0^T \int_{[0,1]} l(\varphi^\epsilon(s,r))\theta(dr)ds &\to \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} \hat{l}(\vartheta_i(y))\rho^i_{\hat{x}(s)}(dy)ds \end{aligned} \tag{5.5.47}$$

in probability, where

$$(\psi^\epsilon(t), \varphi^\epsilon(t,\cdot)) = (\psi^\epsilon_i(t), \varphi^\epsilon_i(t,\cdot)),\ t \in [t_i, t_{i+1}),\ i = 0, 1, \cdots k-1.$$

Proof of the above statement is given in the appendix. Next note that

$$\begin{aligned} &\sum_{i=0}^{k-1} \frac{1}{2}\int_{t_i}^{t_{i+1}} \left(\int_{\mathbb{L}} |u_i(y)|^2 \rho^i_{\hat{x}(s)}(dy) + \frac{|r_i|^2}{\sigma^2} \right) ds & (5.5.48) \\ &\quad = \sum_{i=0}^{k-1} \frac{1}{2}\int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} |u_i(y)|^2 (\rho^i_{\hat{x}(s)}(dy) - \rho^i_{\xi^*(t_i)}(dy))ds \\ &\quad\quad + \sum_{i=0}^{k-1} \frac{1}{2}\int_{t_i}^{t_{i+1}} \left(\int_{\mathbb{L}} |u_i(y)|^2 \varpi^*_i(dy) + \frac{|r_i|^2}{\sigma^2} \right) ds \\ &\quad \leq \frac{\epsilon_0}{2} + M_2 s(\sigma) + \epsilon_0 A_2 \\ &\quad\quad + \sum_{i=0}^{k-1} \frac{1}{2} \left(\int_{\mathbb{L}} |u_i(y)|^2 \varpi^*_i(dy) + \frac{|r_i|^2}{\sigma^2} \right) (t_{i+1} - t_i), & (5.5.49) \end{aligned}$$

where $M_2 = \frac{\beta(L_0,\delta)}{\delta}(2M_0+1)\exp(M_1)$, $A_2 = B(K_1)TM_2$ and the last inequality follows from our choice of d_0 on observing from (5.5.37), (5.5.46), (5.5.28) and (5.5.30) that

$$\begin{aligned} &\sum_{i=0}^{k-1} \frac{1}{2}\int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} |u_i(y)|^2 (\rho^i_{\hat{x}(s)}(dy) - \rho^i_{\xi^*(t_i)}(dy))ds \\ &\quad \leq \frac{\beta(L_0,\delta)}{\delta} \left(\sum_{i=0}^{k-1} m_i \Lambda_i \right) (2d_0 + s(\sigma)\exp(M_1) + \epsilon_0 A_1) \\ &\quad \leq \frac{\beta(L_0,\delta)}{\delta} (2M_0 + 1)(2d_0 + s(\sigma)\exp(M_1) + \epsilon_0 A_1) \\ &\quad \leq \frac{\epsilon_0}{2} + M_2 s(\sigma) + \epsilon_0 A_2. \end{aligned}$$

Similarly

$$
\begin{aligned}
&\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\int_{\mathbb{L}}\hat{l}(\vartheta_i(y))\rho^i_{\hat{x}(s)}(dy)ds\\
&\qquad=\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\int_{\mathbb{L}}\hat{l}(\vartheta_i(y))(\rho^i_{\hat{x}(s)}(dy)-\rho^i_{\xi^*(t_i)}(dy))ds\\
&\qquad\quad+\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\int_{\mathbb{L}}\hat{l}(\vartheta_i(y))\rho^i_{\xi^*(t_i)}(dy)ds\\
&\qquad\leq\frac{\beta(L_0,\delta)}{\delta}\left(\sum_{i=0}^{k-1}m_i\Lambda_i\right)(2d_0+s(\sigma)\exp(M_1))\\
&\qquad\quad+\sum_{i=0}^{k-1}\int_{t_i}^{t_{i+1}}\int_{\mathbb{L}}\hat{l}(\vartheta_i(y))\varpi_i^*(dy)ds\\
&\qquad\leq\frac{\epsilon_0}{2}+M_2s(\sigma)+\epsilon_0A_2+\sum_{i=0}^{k-1}\int_{\mathbb{L}}\hat{l}(\vartheta_i(y))\varpi_i^*(dy)(t_{i+1}-t_i). \qquad (5.5.50)
\end{aligned}
$$

Let

$$
\boldsymbol{u}^\epsilon(t)=(\psi^\epsilon(t),\varphi^\epsilon(t,\cdot)). \qquad (5.5.51)
$$

Define $Z^\epsilon_\sigma=X^\epsilon+U^\epsilon_\sigma$, where $U^\epsilon_\sigma=\sqrt{\epsilon}\sigma B$. Let $\Theta=\max\{2||F||_\infty,M_F\}$, where M_F is as in (5.5.25). Then

$$
\begin{aligned}
&\epsilon\log\mathbb{E}_x\left[\exp\left(-\frac{1}{\epsilon}F(Z^\epsilon_\sigma)\right)\right]\\
&\qquad\leq\epsilon\log\mathbb{E}_x\left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right]+\epsilon\log\mathbb{E}_x\left[\exp\left(\frac{1}{\epsilon}\Theta(||U^\epsilon_\sigma||_T\wedge 1)\right)\right].
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\limsup_{\epsilon\to 0}-\epsilon\log\mathbb{E}_x\left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right]\\
&\qquad\leq\limsup_{\sigma\to 0}\limsup_{\epsilon\to 0}-\epsilon\log\mathbb{E}_x\left[\exp\left(-\frac{1}{\epsilon}F(Z^\epsilon_\sigma)\right)\right]\\
&\qquad\quad+\limsup_{\sigma\to 0}\limsup_{\epsilon\to 0}\epsilon\log\mathbb{E}_x\left[\exp\left(\frac{1}{\epsilon}\Theta(||U^\epsilon_\sigma||_T\wedge 1)\right)\right].
\end{aligned}
$$

Note that, for every $\sigma>0$, as $\epsilon\to 0$, U^ϵ_σ satisfies the Laplace principle with rate function

$$\mathbb{I}_{0,\sigma}(\varphi) = \begin{cases} \frac{1}{2\sigma^2}\int_0^T |\dot{\varphi}(s)|^2 ds & \text{if } \int_0^T |\dot{\varphi}(s)|^2 ds < \infty, \\ \infty & \text{otherwise.} \end{cases} \tag{5.5.52}$$

A straightforward calculation now shows that

$$\limsup_{\sigma\to 0}\limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}_x \left[\exp\left(\frac{1}{\epsilon}\Theta(||U_\sigma^\epsilon||_T \wedge 1)\right)\right] \le \limsup_{\sigma\to 0} \frac{\Theta^2\sigma^2}{2} T = 0.$$

Thus

$$\limsup_{\epsilon\to 0} -\epsilon \log \mathbb{E}_x \left[\exp\left(-\frac{1}{\epsilon}F(X^\epsilon)\right)\right] \le \limsup_{\sigma\to 0}\limsup_{\epsilon\to 0} -\epsilon \log \mathbb{E}_x \left[\exp\left(-\frac{1}{\epsilon}F(Z_\sigma^\epsilon)\right)\right]. \tag{5.5.53}$$

Now consider the expression on the right side of the above display. For $(\psi,\varphi) \in \mathcal{U}$ and $\gamma \in \mathcal{P}_2$ let

$$Z_\epsilon(\psi,\varphi,\gamma) = \mathcal{G}^\epsilon\left(\sqrt{\epsilon}W + \int_0^\cdot \psi(s)ds,\, \epsilon N^{\epsilon^{-1}\varphi}\right) + \sqrt{\epsilon}\sigma B + \int_0^\cdot \gamma(s)ds$$

where $\mathcal{G}^\epsilon$ is as in Section 5.4. Then by the variational representation from [18] (See Theorem 3.1 therein) we have

$$\begin{aligned} &-\epsilon \log \mathbb{E}_x \left[\exp\left(-\frac{1}{\epsilon}F(Z_\sigma^\epsilon)\right)\right] \\ &\qquad = \inf_{\boldsymbol{u}=(\psi,\varphi)\in\mathcal{U}} \inf_{\gamma\in\mathcal{P}_2} \mathbb{E}_x \left[\bar{L}_T(\boldsymbol{u}) + \frac{1}{2\sigma^2}\int_0^T |\gamma(s)|^2 ds + F(Z_\epsilon(\psi,\varphi,\gamma))\right] \\ &\qquad \le \mathbb{E}_x \left[\bar{L}_T(\boldsymbol{u}^\epsilon) + \frac{1}{2\sigma^2}\int_0^T |\zeta(s)|^2 ds + F(Z_\epsilon(\psi^\epsilon,\varphi^\epsilon,\zeta))\right], \end{aligned}$$

where $\boldsymbol{u}^\epsilon = (\psi^\epsilon,\varphi^\epsilon)$ and ζ are as in (5.5.51) and (5.5.42), respectively. From (5.5.47), (5.5.49), (5.5.50), (5.5.36) and (5.5.28) it follows that

$$\begin{aligned}
&\limsup_{\epsilon\to 0}\mathbb{E}_x\left(\bar{L}_T(\boldsymbol{u}^\epsilon) + \frac{1}{2\sigma^2}\int_0^T |\zeta(s)|^2 ds\right) \le \epsilon_0 + 2M_2 s(\sigma) + 2\epsilon_0 A_2\\
&\quad + \sum_{i=0}^{k-1}\left(\int_{\mathbb{L}}\left(\frac{1}{2}|u_i(y)|^2 + \hat{l}(\vartheta_i(y))\right)\varpi_i^*(dy) + \frac{|r_i|^2}{2\sigma^2}\right)(t_{i+1}-t_i)\\
&\quad \le \epsilon_0 + 2M_2 s(\sigma) + 2\epsilon_0 A_2 + (1+\epsilon_0)\sum_{i=0}^{k-1} L_\sigma(\xi^*(t_j),\dot{\xi}^*(t_j))(t_{i+1}-t_i) + \epsilon_0 T\\
&\quad \le (1+T+M_0)\epsilon_0 + 2M_2 s(\sigma) + 2\epsilon_0 A_2 + \sum_{i=0}^{k-1} L_\sigma(\xi^*(t_j),\dot{\xi}^*(t_j))(t_{i+1}-t_i).
\end{aligned} \tag{5.5.54}$$

An application of Girsanov's theorem shows that $Z_\epsilon(\psi^\epsilon,\varphi^\epsilon,\zeta)) = \tilde{Z}^\epsilon$ a.s. and as noted below (5.5.41), $Z_\epsilon(\psi^\epsilon,\varphi^\epsilon,\zeta)) = \tilde{Z}^\epsilon$ converges in probability to z. Thus, using (5.5.45), we have

$$\limsup_{\epsilon\to 0}\mathbb{E}_x\left[F(Z_\epsilon(\psi^\epsilon,\varphi^\epsilon,\beta)\right] \le F(\xi^*) + M_F|\zeta|_T + M_F\left(\frac{\epsilon_0}{2} + s(\sigma)\exp(M_1) + \epsilon_0 A_1\right). \tag{5.5.55}$$

Combining (5.5.54), (5.5.43), (5.5.55) and (5.5.27) we have

$$\begin{aligned}
&\limsup_{\epsilon\to 0} -\epsilon\log\mathbb{E}_x\left[\exp\left(-\frac{1}{\epsilon}F(Z_\sigma^\epsilon)\right)\right]\\
&\qquad \le (1+T+M_0)\epsilon_0 + 2M_2 s(\sigma) + 2\epsilon_0 A_2\\
&\qquad\quad + \sum_{i=0}^{k-1} L_\sigma(\xi^*(t_j),\dot{\xi}^*(t_j))(t_{i+1}-t_i) + F(\xi^*)\\
&\qquad\quad + M_F s(\sigma) + M_F\left(\frac{\epsilon_0}{2} + s(\sigma)\exp(M_1) + \epsilon_0 A_1\right)\\
&\qquad \le \inf_{\xi\in\mathbb{U}}\{F(\xi) + \mathbb{I}(\xi)\} + (2+T+M_0+M_F+2A_2+A_1)\epsilon_0\\
&\qquad\quad + (M_F(1+\exp(M_1)) + 2M_2)s(\sigma).
\end{aligned}$$

Note that the constants M_F, M_1, M_2, A_1, A_2 depend only on F, the coefficients a, b, and $\bar{M}$ (and hence on ϵ_0), but not on σ. The result now follows on sending first $\sigma\to 0$ and then $\epsilon_0\to 0$. $\square$

5.6 Appendix.

Proof of Lemma 5.2.1.

In view of Assumption 5.2.1 (2), it suffices to show that ρ_x is Lipschitz continuous in x. The $\mathbb{L}$ valued Markov chain with transition probabilities $p^x_{yz} = \sum_{n=1}^{\ell} \hat{R}_n(x, y, z)$, $(y, z) \in \mathbb{L}$, is ergodic for each $x \in \mathbb{R}^d$. Denote the unique invariant measure of this chain by π^x. From (5.2.1) it follows that p^x_{yz} is Lipschitz in x and $\inf_{y,z\in\mathbb{L}} \inf_{x\in\mathbb{R}^d} p^x_{yz} > 0$. From Lemma 3.1 in [42], π^x is given as a ratio of polynomials in $\{p^x_{yz}\}_{y,z\in\mathbb{L}}$. Thus $x \mapsto \pi^x(y)$ is Lipschitz for every $y \in \mathbb{L}$. The result now follows on noting that $\rho_x(y) = \frac{\pi_x(y)}{c(x,y)}$ along with the observation that $x \mapsto c(x, y)$ is Lipschitz for every $y \in \mathbb{L}$ (from (5.2.1)) and the lower bound on $c(x, y)$. $\square$

Proof of (5.3.5) and (5.3.6).

Denote the set $E_{x,y} \Delta E_{x',y}$ by $\bar{E}$. Then

$$\begin{aligned}|\hat{c}^\eta(x, y) - \hat{c}^\eta(x', y)| =& |\eta(E_{x,y}) - \eta(E_{x',y})| \le \eta(\bar{E}) \\ =& \int_{\bar{E}} 1 \cdot \frac{d\eta}{d\theta}(r)\theta(dr) \le \int_{\bar{E}} e^m \theta(dr) + \frac{1}{m}\int_{\bar{E}} l(\frac{d\eta}{d\theta}(r))\theta(dr) \\ \le& e^m \kappa_2 |x - x'| + \frac{1}{m}\hat{l}(\eta),\end{aligned}$$

where the inequality on the second line follows from (5.3.4) and the last inequality follows from (5.2.1).

Similarly,

$$\begin{aligned}&\sup_{y'\in\mathbb{L}, y'\neq y} \left| \hat{c}^\eta(x, y)\hat{R}^\eta(x, y, y') - \hat{c}^\eta(x', y)\hat{R}^\eta(x', y, y') \right| \\ &\qquad = \sup_{y'\in\mathbb{L}, y'\neq y} |\eta(E^{y'}_{x,y}) - \eta(E^{y'}_{x',y})| \le e^m \kappa_2 |x - x'| + \frac{1}{m}\hat{l}(\eta).\end{aligned}$$

Combining the above estimates we have (5.3.5). Proof of (5.3.6) is now immediate on observing that

$$|\Pi_x^\eta \phi(y) - \Pi_{x'}^\eta \phi(y)| \leq |\phi|_\infty \left|\hat{c}^\eta(x,y) - \hat{c}^\eta(x',y)\right| + |\phi|_\infty \sup_{y' \in \mathbb{L}, y' \neq y} \left|\hat{c}^\eta(x,y)\hat{R}^\eta(x,y,y') - \hat{c}^\eta(x',y)\hat{R}^\eta(x',y,y')\right|.$$

□

Proof of (5.4.12).

Changing the order of the integration, we can rewrite the following term as

$$\begin{aligned}
&\int_0^t \frac{1}{\Delta_\epsilon}\int_s^{(s+\Delta_\epsilon)\wedge T} a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)duds \\
&= \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_0^u a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu \\
&\quad + \int_{\Delta_\epsilon}^t \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu \\
&\quad + \int_t^{(t+\Delta_\epsilon)\wedge T} \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu.
\end{aligned}$$

Also, we have

$$\int_0^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))\psi^\epsilon(s)ds = \int_0^t \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u a(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu.$$

Thus

$$\begin{aligned}
&\left|\int_0^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))\psi^\epsilon(s)ds - \int_0^t \frac{1}{\Delta_\epsilon}\int_s^{s+\Delta_\epsilon} a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)duds\right| \\
&\leq \left|\int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u a(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu - \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon}\int_0^u a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu\right| \\
&+ \left|\int_{\Delta_\epsilon}^t \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u a(\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu - \int_{\Delta_\epsilon}^t \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^u a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu\right| \\
&+ \left|\int_t^{(t+\Delta_\epsilon)\wedge T} \frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu\right| \\
&= T_t^{(1)} + T_t^{(2)} + T_t^{(3)}
\end{aligned}$$

For $T^{(1)}$ we have

$$
\begin{aligned}
(T_t^{(1)})^2 &\le 2\left|\int_0^{\Delta_\epsilon}\frac{1}{\Delta_\epsilon}\int_{u-\Delta_\epsilon}^{u} a(\tilde{X}^\epsilon(u),\tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu\right|^2 \\
&\quad + 2\left|\int_0^{\Delta_\epsilon}\frac{1}{\Delta_\epsilon}\int_0^{u} a(\tilde{X}^\epsilon(s),\tilde{Y}^\epsilon(u))\psi^\epsilon(u)dsdu\right|^2 \\
&\le 4\left(\kappa_1\sup_{0\le s\le t}(|\tilde{X}^\epsilon(s)|+1)\right)^2\int_0^{\Delta_\epsilon}|\psi^\epsilon(u)|du \\
&\le 8\kappa_1^2(\sup_{0\le s\le t}|\tilde{X}^\epsilon(s)|^2+1)\cdot\Delta_\epsilon 2M.
\end{aligned}
$$

Similarly,

$$
(T_t^{(3)})^2 \le 4\kappa_1^2(\sup_{0\le s\le t}|\tilde{X}^\epsilon(s)|^2+1)\cdot\Delta_\epsilon 2M.
$$

Thus in view of (5.4.8), $\sup_{0\le t\le T}\mathbb{E}(T_t^{(i)})^2\to 0$, as $\epsilon\to 0$, for $i=1,3$.

For the second term, using the Lipschitz property of a, we have,

$$
\begin{aligned}
(T_t^{(2)})^2 &\le \int_{\Delta_\epsilon}^t\int_{u-\Delta_\epsilon}^u\frac{1}{\Delta_\epsilon^2}d_{lip}^2|\tilde{X}^\epsilon(u)-\tilde{X}^\epsilon(s)|^2dsdu\cdot\int_{\Delta_\epsilon}^t\int_{u-\Delta_\epsilon}^u(\psi^\epsilon(u))^2dsdu \\
&\le \int_{\Delta_\epsilon}^t\int_{u-\Delta_\epsilon}^u\frac{1}{\Delta_\epsilon^2}d_{lip}^2|\tilde{X}^\epsilon(u)-\tilde{X}^\epsilon(s)|^2dsdu\cdot\Delta_\epsilon 2M,
\end{aligned}
$$

where d_{lip} is as in (5.5.26). Using (5.4.9) we have

$$
\begin{aligned}
\sup_{0\le t\le T}\mathbb{E}(T_t^{(2)})^2 &\le \frac{d_{lip}^2}{\Delta_\epsilon^2}T\Delta_\epsilon\sup_{\Delta_\epsilon\le t\le T}\mathbf{E}\left(\sup_{u-\Delta_\epsilon\le s\le u}|\tilde{X}^\epsilon(u)-\tilde{X}^\epsilon(s)|^2\right)\Delta_\epsilon 2M \\
&\le d_{lip}^2c_1T\Delta_\epsilon 2M,
\end{aligned}
$$

which converges to 0 as $\epsilon\to 0$. Combining the above estimates, we have (5.4.12). $\square$

Proof of (5.4.16).

Recall that

$$
\bar{L}_T(u^\epsilon)=\frac{1}{2}\int_0^T|\psi^\epsilon(s)|^2ds+\int_{[0,T]\times[0,1]}l(\varphi^\epsilon(r,s))\theta(dr)ds,
$$

and

$$
\begin{aligned}
&\int_{\mathbb{H}_T}\left[\frac{1}{2}|z|^2+\hat{l}(\eta)\right]Q^\epsilon(d\mathbf{v}) \\
&\quad=\frac{1}{2}\int_0^T\frac{1}{\Delta_\epsilon}\int_s^{s+\Delta_\epsilon}|\psi^\epsilon(u)|^2duds+\int_0^T\frac{1}{\Delta_\epsilon}\int_s^{s+\Delta_\epsilon}\hat{l}(\eta^\epsilon(u))duds,
\end{aligned}
$$

where η^ϵ is as in (5.4.13).

Similar to the proof of (5.4.12), changing the order of integration, we can rewrite the term $\int_{\mathbb{H}_T} |z|^2 Q^\epsilon(d\mathbf{v})$ as

$$\int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon} \int_0^u |\psi^\epsilon(u)|^2 dsdu + \int_{\Delta_\epsilon}^T \frac{1}{\Delta_\epsilon} \int_{u-\Delta_\epsilon}^u |\psi^\epsilon(u)|^2 dsdu + \int_T^{T+\Delta_\epsilon} \frac{1}{\Delta_\epsilon} \int_{u-\Delta_\epsilon}^T |\psi^\epsilon(u)|^2 dsdu,$$

where the third term is 0 using our convention that $\psi^\epsilon(u) = 0$, for $u > T$.

Also, we have

$$\int_0^T |\psi^\epsilon(s)|^2 ds = \int_0^T \frac{1}{\Delta_\epsilon} \int_{u-\Delta_\epsilon}^u |\psi^\epsilon(u)|^2 dsdu.$$

Then

$$\begin{aligned} \int_0^T |\psi^\epsilon(s)|^2 ds - \int_{\mathbb{H}_T} |z|^2 Q^\epsilon(d\mathbf{v}) \\ &= \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon} \int_0^{\Delta_\epsilon} |\psi^\epsilon(u)|^2 dsdu - \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon} \int_0^u |\psi^\epsilon(u)|^2 dsdu \\ &= \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon} \int_u^{\Delta_\epsilon} |\psi^\epsilon(u)|^2 dsdu \geq 0. \end{aligned}$$

For the second term we have in a similar manner

$$\int_{[0,T]\times[0,1]} l(\varphi^\epsilon(r,s))\theta(dr)ds - \int_{\mathbb{H}_T} \hat{l}(\eta) Q^\epsilon(d\mathbf{v}) = \int_0^{\Delta_\epsilon} \frac{1}{\Delta_\epsilon} \int_u^{\Delta_\epsilon} \hat{l}(\eta^\epsilon(u)) dsdu \geq 0.$$

The result follows. $\square$

Proof of (5.4.5).

Recall from the discussion below (5.4.1) that $X^\epsilon = \mathcal{G}^\epsilon(\sqrt{\epsilon}W, \epsilon N^{\epsilon^{-1}})$ is a solution of (5.2.3)-(5.2.4). Let $\bar{\varphi}^\epsilon(r,s) = (\varphi^\epsilon(r,s))^{-1}$, $(r,s) \in \mathbb{R} \times [0,1]$ and let

$$\mathcal{E}_t^\epsilon = \exp\left\{ \int_{[0,1]\times[0,t]} \log(\bar{\varphi}^\epsilon(r,s)) N^{1/\epsilon}(dr \times ds) + \int_{[0,1]\times[0,t]} (1 - \bar{\varphi}^\epsilon(r,s))\theta(dr)ds \right\}$$

and

$$\bar{\mathcal{E}}_t^\epsilon = \exp\left\{ \int_0^t \langle \psi^\epsilon(s), dW(s) \rangle - \frac{1}{2} \int_0^t |\psi^\epsilon(s)|^2 ds \right\}.$$

Define $\tilde{\mathcal{E}}_t^\epsilon = \mathcal{E}_t^\epsilon \bar{\mathcal{E}}_t^\epsilon$. Then $\tilde{\mathcal{E}}_t^\epsilon$ is a $\{\mathcal{F}_t\}$ martingale and consequently, $\mathbb{Q}_T^\epsilon(G) = \int_G \tilde{\mathcal{E}}_T^\epsilon d\mathbb{P}$, $G \in \mathcal{F}$, defines a probability measure on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\})$ and furthermore $\mathbb{P}$ and $\mathbb{Q}_T^\epsilon$ are mutually absolutely continuous. Also, under $\mathbb{Q}_T^\epsilon$, $(\sqrt{\epsilon}W + \int_0^\cdot \psi^\epsilon(s)ds, \epsilon N^{\epsilon^{-1}\varphi^\epsilon})$ has the same law as that of $(\sqrt{\epsilon}W, \epsilon N^{\epsilon^{-1}})$ under $\mathbb{P}$. The statement in (5.4.5) is now immediate. □

Proof of Lemma 5.5.6.

The $\mathbb{L}$ valued Markov chain with transition probabilities $\hat{p}_{yz}^x = \sum_{n=1}^{\ell} \hat{R}_n^{\vartheta^*(y,\cdot)}(x, y, z)$, $(y, z) \in \mathbb{L}$, is ergodic for each $x \in \mathbb{R}^d$. Denote the unique invariant measure of this chain by $\hat{\pi}^x$. From (5.2.1) and (5.5.18) it follows that $\hat{p}_{yz}^x$ is Lipschitz in x and $\inf_{y,z\in\mathbb{L}} \inf_{x\in\mathbb{R}^d} \hat{p}_{yz}^x > 0$. As in Appendix 5.6, we now have that $x \mapsto \hat{\pi}^x(y)$ is Lipschitz for every $y \in \mathbb{L}$. The result now follows on noting that $\varpi_x(y) = \frac{\hat{\pi}_x(y)}{\hat{c}^{\vartheta^*(y,\cdot)}(x,y)}$ along with the observation that $x \mapsto \hat{c}^{\vartheta^*(y,\cdot)}(x, y)$ is Lipschitz for every $y \in \mathbb{L}$ (from (5.2.1) and (5.5.18)) and the lower bound on $\hat{c}^{\vartheta^*(y,\cdot)}(x, y)$ in (5.5.12). □

Proof of (5.5.47).

It suffices to show that for each $i = 0, \cdots, k-1$,

$$\begin{aligned} \frac{1}{2}\int_{t_i}^{t_{i+1}} |\psi_i^\epsilon(s)|^2 ds \to& \frac{1}{2}\int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} |u_i(y)|^2 \rho_{\hat{x}(s)}^i(dy)ds, \\ \int_{t_i}^{t_{i+1}} \int_{[0,1]} l(\varphi_i^\epsilon(s, r))\theta(dr)ds \to& \int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} \hat{l}(\vartheta_i(y))\rho_{\hat{x}(s)}^i(dy)ds \end{aligned}$$

in probability. Recalling the definition of ψ_i^ϵ and φ_i^ϵ in (5.5.40) we have that to prove the above convergences, it is enough to show that for any function $g : \mathbb{L} \to \mathbb{R}$, we have

$$\int_{t_i}^{t_{i+1}} g(\tilde{Y}^\epsilon(s))ds \to \int_{t_i}^{t_{i+1}} \int_{\mathbb{L}} g(y)\rho_{\hat{x}(s)}^i(dy)ds$$

in probability. We only consider the case $i = 0$ and write $t_1 = T$. Then for $t \in [0, T]$, $\tilde{Y}^\epsilon$ is given through the system of equations

$$\begin{aligned}
\tilde{X}^\epsilon(t) =& \tilde{X}^\epsilon(0) + \int_0^t b(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))ds + \sqrt{\epsilon}\int_0^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))dW(s) \\
&+ \int_0^t a(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s))u_0(\tilde{Y}^\epsilon(s))ds \\
\tilde{Y}^\epsilon(t) =& \tilde{Y}^\epsilon(t_j) + \int_{t_j}^t \int_{r\in[0,1]} k(\tilde{X}^\epsilon(s), \tilde{Y}^\epsilon(s-), r)N^{\frac{1}{\epsilon}\varphi_0^*(\tilde{Y}^\epsilon(s-))}(dr \times ds).
\end{aligned} \tag{5.6.1}$$

Note that

$$\begin{aligned}
\mathbb{P}&\left(\left|\int_0^t g(\tilde{Y}^\epsilon(s))ds \; - \; \int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\hat{x}(s)}(dy)ds\right| > \delta\right) \\
&\qquad\le \mathbb{P}\left(\left|\int_0^t g(\tilde{Y}^\epsilon(s))ds - \int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(s)}(dy)ds\right| > \frac{\delta}{2}\right) \\
&\qquad+ \mathbb{P}\left(\left|\int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(s)}(dy)ds - \int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\hat{x}(s)}(dy)ds\right| > \frac{\delta}{2}\right).
\end{aligned}$$

Since $\tilde{X}^\epsilon$ converges in probability to $\hat{x}$, with the uniform metric on $\mathbb{U}$ (see below (5.5.40)), we have from (5.5.38) that

$$\begin{aligned}
\mathbb{P}&\left(\left|\int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(s)}(dy)ds \; - \; \int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\hat{x}(s)}(dy)ds\right| > \frac{\delta}{2}\right) \\
&\qquad\le \mathbb{P}\left(||g||_\infty \beta(L_0, \delta) \sup_{0\le s\le t} ||\tilde{X}^\epsilon(s) - \hat{x}(s)||t > \frac{\delta}{2}\right) \to 0.
\end{aligned}$$

Thus we only need to show, for all $t \in [0T]$,

$$\mathbf{E}\left(\left|\int_0^t g(\tilde{Y}^\epsilon(s))ds - \int_0^t \int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(s)}(dy)ds\right|\right) \to 0, \text{ as } \epsilon \to 0. \tag{5.6.2}$$

Fix a sequence $\{h_\epsilon\}_{\epsilon>0}$ such that, as $\epsilon \to 0$,

$$h_\epsilon \to 0, \quad \frac{\epsilon}{h_\epsilon} \to 0 \text{ and } \frac{h_\epsilon^{3/2}}{\epsilon} \to 0.$$

By a change of order of integration we have,

$$
\begin{aligned}
&\left|\int_0^t g(\tilde{Y}^\epsilon(s))ds - \int_0^t \left(\frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\tilde{Y}^\epsilon(s))ds\right)du\right| \\
&\qquad\qquad \leq \left|\int_0^{h_\epsilon}\left(\frac{1}{h_\epsilon}\int_0^s g(\tilde{Y}^\epsilon(s))du\right)ds\right| + \left|\int_0^{h_\epsilon}\left(\frac{1}{h_\epsilon}\int_{s-h_\epsilon}^s g(\tilde{Y}^\epsilon(s))du\right)ds\right| \\
&\qquad\qquad\quad + \left|\int_t^{t+h_\epsilon}\left(\frac{1}{h_\epsilon}\int_{s-h_\epsilon}^t g(\tilde{Y}^\epsilon(s))du\right)ds\right| \\
&\qquad\qquad \leq 3||g||h_\epsilon
\end{aligned}
\tag{5.6.3}
$$

Also, for every $u \in [0, T]$, for any $\epsilon_0 > 0$, there exists a compact set K, such that, $\mathbb{P}(||\tilde{X}^\epsilon(u)|| \in K^c) \leq \epsilon_0 > 0$. Following arguments along the lines in Theorem II.7 in [73], we have,

$$
\mathbf{E}\left(\mathbf{E}_{\tilde{X}^\epsilon(u),\tilde{Y}^\epsilon(u)}\left|\frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\tilde{Y}^\epsilon(s))ds - \frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\bar{Y}_u^\epsilon(s))ds\right|\right) \leq \kappa h_\epsilon^{3/2}||g||/\epsilon + 2\epsilon_0||g||,
\tag{5.6.4}
$$

where $\mathbf{E}_{\tilde{X}^\epsilon(u),\tilde{Y}^\epsilon(u)}$ denotes the conditional expectation given $\{\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u)\}$, κ depends only on K, and $\{\bar{Y}_u^\epsilon(s)\}_{s\geq u}$ is the solution of the system of equations

$$
\begin{aligned}
\bar{X}_u^\epsilon(t) =& \tilde{X}^\epsilon(u) + \int_u^t b(\bar{X}_u^\epsilon(s), \bar{Y}_u^\epsilon(s))ds + \sqrt{\epsilon}\int_u^t a(\bar{X}_u^\epsilon(s), \bar{Y}_u^\epsilon(s))dW(s) \\
&+ \int_u^t a(\bar{X}_u^\epsilon(s), \bar{Y}_u^\epsilon(s))u_0(\bar{Y}_u^\epsilon(s))ds, \\
\bar{Y}_u^\epsilon(t) =& \tilde{Y}^\epsilon(u) + \int_u^t \int_{r\in[0,1]} k(\tilde{X}^\epsilon(u), \bar{Y}_u^\epsilon(s-), r)N^{\frac{1}{\epsilon}\varphi^*(\bar{Y}_u^\epsilon(s-))}(dr \times ds).
\end{aligned}
$$

In view of (5.6.3) and (5.6.4), to show (5.6.2), it is enough to show that, for all $t \in [0, T]$,

$$
\mathbf{E}\left|\int_0^t \frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\bar{Y}_u^\epsilon(s))dsdu - \int_0^t\int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(s)}(dy)ds\right| \to 0.
\tag{5.6.5}
$$

Next note that

$$
\begin{aligned}
&\mathbf{E}\left|\int_0^t \frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\bar{Y}_u^\epsilon(s))dsdu - \int_0^t\int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(s)}(dy)ds\right| \\
&\qquad\qquad \leq \int_0^t \mathbf{E}\left|\frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\bar{Y}_u^\epsilon(s))ds - \int_{\mathbb{L}} g(y)\rho^0_{\tilde{X}^\epsilon(u)}(dy)\right|du.
\end{aligned}
\tag{5.6.6}
$$

For every u fixed,

$$\frac{1}{h_\epsilon}\int_u^{u+h_\epsilon} g(\bar{Y}_u^\epsilon(s))ds = \frac{\epsilon}{h_\epsilon}\int_0^{h_\epsilon/\epsilon} g(\bar{Y}_u^\epsilon(u+\epsilon s))ds.$$

Note that, conditional on $\{\tilde{X}^\epsilon(u), \tilde{Y}^\epsilon(u)\}$, $\{\bar{Y}_u^\epsilon(u+\epsilon s)\}_{s\geq 0}$ is a $\mathbb{L}$ valued ergodic Markov process with generator $\Pi_{\tilde{X}^\epsilon(u)}^{\vartheta_0}$ whose unique invariant measure is $\rho_{\tilde{X}^\epsilon(u)}^0$. Thus by the ergodic theorem, and recalling that $h_\epsilon/\epsilon \to \infty$ as $\epsilon \to 0$, we have

$$\mathbf{E}_{\tilde{X}^\epsilon(u),\tilde{Y}^\epsilon(u)}\left|\frac{\epsilon}{h_\epsilon}\int_u^{u+h_\epsilon/\epsilon} g(\bar{Y}_u^\epsilon(\epsilon s))ds - \int_{\mathbb{L}} g(y)\rho_{\tilde{X}^\epsilon(u)}^0(dy)\right| \to 0, \text{ as } \epsilon \to 0.$$

The convergence in (5.6.5) now follows on using the dominated convergence theorem.

□

Chapter 6

A Numerical Scheme for Invariant Distributions of Constrained Diffusions

6.1 Introduction.

Reflected diffusion processes in polyhedral domains have been proposed as approximate models for critically loaded stochastic processing networks. Starting with the influential paper of Reiman[66], there have been many works[65, 27, 62, 85, 55, 84] that justify approximations via reflected diffusions rigorously by establishing a limit theorem under appropriate heavy traffic assumptions. Many performance measures for stochastic networks are formulated to capture the long term behavior of the system and a key object involved in the computation of such measures is the corresponding steady state distribution. Although classical heavy traffic limit theorems only justify approximations of the network behavior through the associated diffusion limit over any fixed finite time horizon, there are now several results[44, 20, 21] that prove, for certain generalized Jackson network models, the convergence of steady state distributions of stochastic networks to those of the associated limit diffusions. Such limit theorems then lead to the important question: How does one compute the stationary distributions of reflected diffusions? Indeed, one of the main motivations for introducing diffusion approximations in the study of stochastic processing systems is the expectation that diffusion models are easier to analyze than their stochastic network counterparts. Classical results of Harrison and Williams [48] show that under certain geometric conditions on the underlying problem data, stationary densities of reflected

Brownian motions have explicit product form expressions. However, once one moves away from this special family of models there are no explicit formulas and thus one needs to use numerical procedures.

The objective of the current work is to propose and study the performance of one such numerical procedure for computing stationary distributions of reflected diffusions in polyhedral domains. For diffusions in $\mathbb{R}^m$ there are two basic approaches for computation of invariant distributions: PDE methods and Monte-Carlo methods. PDE approaches are based on the well known basic property that invariant densities of diffusions can be characterized as solutions of certain stationary Fokker-Planck equations. For reflected Brownian motions in polyhedral domains the papers[26, 54, 25] develop similar characterization results. The characterization in this case is formulated for the invariant density together with certain boundary densities and is given in terms of the second order differential operator describing the underlying unconstrained dynamics and a collection of first order operators corresponding to the boundary reflections. Using this characterization as a starting point Dai and Harrison[25] develop an approximation scheme for the stationary density by constructing projections on to certain finite dimensional Hilbert spaces that are described in terms of the above collection of differential operators. Although PDE methods such as above are quite efficient for settings where the state dimension m is small, one finds that Monte-Carlo methods, based on the use of the ergodic theorem, have advantages in higher dimensions. With this in mind, we will propose and study here a Monte-Carlo method for the computation of stationary distributions. Approximations of invariant distributions of diffusions in $\mathbb{R}^m$ using simulation of paths have been studied in several works [4, 64, 76, 75, 57]. One of the key difficulties in using simulation methods is that paths of diffusions cannot be simulated exactly and so one has to contend with two sources of errors: Discretization of the SDE and finite time empirical average approximation for the steady state behavior. In particular, the long term behavior of the discretized SDE could, in general, be quite different from that of the original system and thus a performance analysis of such Monte-Carlo schemes requires a careful understanding of the stability properties of the underlying systems.

The Monte-Carlo approach studied in the current work is inspired by the papers [4], [64], [57] which have analyzed the properties of weighted empirical measures constructed from a Euler scheme, based on a single sequence of time discretization steps decreasing to zero, for diffusions in $\mathbb{R}^m$. For multi-dimensional diffusions with reflection one first needs to describe a suitable analog of an ‘Euler discretization step’. In order to do so, we begin with a precise description of the stochastic dynamical system of interest.

Let $G \subset \mathbb{R}^m$ be the convex polyhedral cone in $\mathbb{R}^m$ with the vertex at origin given as the intersection of half spaces G_i, $i = 1, \ldots, N$. Let n_i be the unit vector associated with G_i via the relation

$$G_i = \{x \in \mathbb{R}^m : \langle x, n_i \rangle \geq 0\}.$$

Denote the boundary of a set $S \subset \mathbb{R}^m$ by ∂S. We will denote the set $\{x \in \partial G : \langle x, n_i \rangle = 0\}$ by F_i. For $x \in \partial G$, define the set, $n(x)$, of unit inward normals to G at x by

$$n(x) \doteq \{r : |r| = 1, \langle r, x - y \rangle \leq 0, \forall y \in G\}.$$

With each face F_i we associate a unit vector d_i such that $\langle d_i, n_i \rangle > 0$. This vector defines the direction of constraint associated with the face F_i. For $x \in \partial G$ define

$$d(x) \doteq \left\{ d \in \mathbb{R}^m : d = \sum_{i \in \mathrm{In}(x)} \alpha_i d_i ; \alpha_i \geq 0; |d| = 1 \right\},$$

where

$$\mathrm{In}(x) \doteq \{i \in \{1, 2, \cdots, N\} : \langle x, n_i \rangle = 0\}.$$

Roughly speaking, the set $d(x)$ represents the set of permissible directions of constraint available at a point $x \in \partial G$. In a typical stochastic network setting this set valued function is determined from the routing structure of the network and governs the precise constraining mechanism that is used. This mechanism specifies how a RCLL trajectory ψ with values in $\mathbb{R}^m$ is constrained to form a new trajectory with values in G, through the associated Skorohod problem, which is defined as follows.

Let $D([0,\infty) : \mathbb{R}^m)$ denote the set of functions mapping $[0,\infty)$ to $\mathbb{R}^m$ that are right continuous and have left limits. We endow $D([0,\infty) : \mathbb{R}^m)$ with the usual Skorokhod topology. Let

$$D_G([0,\infty) : \mathbb{R}^m) \doteq \{\psi \in D([0,\infty) : \mathbb{R}^m) : \psi(0) \in G\}.$$

For $\eta \in D([0,\infty) : \mathbb{R}^m)$ let $|\eta|(T)$ denote the total variation of η on $[0,T]$ with respect to the Euclidean norm on $\mathbb{R}^m$.

Definition 6.1.1. Let $\psi \in D_G([0,\infty) : \mathbb{R}^m)$ be given. Then the pair $(\phi, \eta) \in D([0,\infty) : \mathbb{R}^m) \times D([0,\infty) : \mathbb{R}^m)$ solves the Skorokhod problem (SP) for ψ with respect to G and d if and only if $\phi(0) = \psi(0)$, and for all $t \in [0,\infty)$

(i) $\phi(t) = \psi(t) + \eta(t)$;

(ii) $\phi(t) \in G$;

(iii) $|\eta|(t) < \infty$;

(iv) $|\eta|(t) = \int_{[0,t]} I_{\{\phi(s)\in\partial G\}} d|\eta|(s)$;

(v) There exists Borel measurable $\gamma : [0,\infty) \to \mathbb{R}^m$ such that $\gamma(t) \in d(\phi(t))$, $d|\eta|$-almost everywhere and
$$\eta(t) = \int_{[0,t]} \gamma(s) d|\eta|(s).$$

In the above definition ϕ represents the constrained version of ψ and η describes the correction applied to ψ in order to produce ϕ. On the domain $D \subset D_G([0,\infty) : \mathbb{R}^m)$ on which there is a unique solutions to the Skorokhod problem we define the Skorokhod map (SM) Γ as $\Gamma(\psi) \doteq \phi$, if $(\phi, \psi - \phi)$ is the unique solution of the Skorokhod problem posed by ψ. We will make the following assumption on the regularity of the Skorokhod map defined by the data $\{(d_i, n_i); i = 1, 2, \cdots, N\}$.

Condition 6.1.1. *The Skorokhod map is well defined on all of $D_G([0,\infty) : \mathbb{R}^m)$, that is, $D = D_G([0,\infty) : \mathbb{R}^m)$ and the SM is Lipschitz continuous in the following sense. There exists a $K < \infty$ such that for all $\phi_1, \phi_2 \in D_G([0,\infty) : \mathbb{R}^m)$,*

$$\sup_{0 \le t < \infty} |\Gamma(\phi_1)(t) - \Gamma(\phi_2)(t)| < K \sup_{0 \le t < \infty} |\phi_1(t) - \phi_2(t)|.$$

We will also make the following assumption on the problem data.

Condition 6.1.2. *For every $x \in \partial G$, there is a $n \in n(x)$ such that $\langle d, n \rangle > 0$ for all $d \in d(x)$.*

The above condition is equivalent to the assumption that the $N \times N$ matrix with $(i,j)^{th}$ entry $\langle d_i, n_j \rangle$ is complete-S (see [34, 67]). When $G = \mathbb{R}_+^m$ and $N = m$, it is known that Condition 6.1.1 implies Condition 6.1.2 (see [77]). An important consequence of Condition 6.1.2 that will be used in our work is the following result from [10] (see also [28]).

Lemma 6.1.1. *Suppose that Condition 6.1.2 holds. Then there exists a $g \in C_b^2(\mathbb{R}^m)$ such that*

$$\langle \nabla g(x), d_i \rangle \ge 1 \quad \forall x \in F_i, \quad i \in \{1, ..., N\}.$$

We remark here that the function constructed in [10] is defined only on G, however a minor modification of the construction there gives a C^2 extension to all of $\mathbb{R}^m$.

We refer the reader to [46, 34, 35] for sufficient conditions under which Condition 6.1.1 and Condition 6.1.2 hold. For example, the paper [35] shows that if $G = \mathbb{R}^m_+$, $N = m$ and the square matrix $D = [d_1, ..., d_m]$ is of the form $D = M(I - V)$, where M is a diagonal matrix with positive diagonal entries, V is off diagonal and the spectral radius of $|V|$ is less than 1, then both Conditions 6.1.1 and 6.1.2 hold. Here $|V|$ represents the matrix with entries $(|V_{ij}|)$, where V_{ij} is the (i, j)-th entry of V.

We now describe the constrained diffusion process that will be studied in this paper. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space on which is given a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual hypotheses. Let $(W(t), \mathcal{F}_t)$ be a m-dimensional standard Wiener process on the above probability space. For $x \in G$, denote by X^x the unique solution to the following stochastic integral equation,

$$X^x(t) = \Gamma\left(x + \int_0^{\cdot} \sigma(X^x(s))dW(s) + \int_0^{\cdot} b(X^x(s))ds\right)(t), \tag{6.1.1}$$

where $\sigma : G \to \mathbb{R}^{m\times m}$ and $b : G \to \mathbb{R}^m$ are maps satisfying the following condition.

Condition 6.1.3. *There exists $a_1 \in (0, \infty)$ such that*

$$|\sigma(x) - \sigma(y)| + |b(x) - b(y)| \leq a_1|x - y| \quad \forall x, y \in G$$

and

$$|\sigma(x)| \leq a_1, \quad |b(x)| \leq a_1, \quad \forall x \in G.$$

Unique solvability of (6.1.1) can be shown using the above condition and the regularity assumption on the Skorokhod map. In fact, the classical method of Picard iteration gives the following:

Theorem 6.1.1. *For each $x \in G$ there exists a unique pair of continuous $\{\mathcal{F}_t\}$ adapted process $(X^x(t), k(t))_{t\geq 0}$ and a progressively measurable process $(\gamma(t))_{t\geq 0}$ such that the following hold:*

(i) $X^x(t) \in G$, *for all* $t \geq 0$, *a.s.*

(ii) *For all* $t \geq 0$,

$$X^x(t) = x + \int_0^t \sigma(X^x(s))dW(s) + \int_0^t b(X^x(s))ds + k(t), \tag{6.1.2}$$

a.s.

(iii) *For all* $T \in [0, \infty)$,

$$|k|(T) < \infty \quad a.s.$$

(iv) *Almost surely, for every* $t \geq 0$,

$$|k|(t) = \int_0^t I_{\{X^x(s) \in \partial G\}} d|k|(s),$$

$k(t) = \int_0^t \gamma(s) d|k|(s)$, *and* $\gamma(s) \in d(X^x(s))$ *a.e,* $[d|k|]$.

In this work we are interested in the invariant distributions of the strong Markov process $\{X^x\}$. One of the basic results due to Harrison and Williams[47] (see also [13]) on invariant distributions of such Markov processes says that if b and σ are constants and σ is invertible, then X^x has a unique invariant probability measure if $b \in \mathcal{C}^o$ (the interior of $\mathcal{C}$), where

$$\mathcal{C} \doteq \left\{ -\sum_{i=1}^N \alpha_i d_i : \alpha_i \geq 0; i \in \{1, \cdots, N\} \right\}.$$

This result was extended to a setting with state dependent coefficients in [1] as follows. We introduce the following two additional assumptions. For $\delta \in (0, \infty)$, define

$$\mathcal{C}(\delta) \doteq \{v \in \mathcal{C} : \text{dist}(v, \partial\mathcal{C}) \geq \delta\}.$$

Condition 6.1.4. *There exists a* $\delta \in (0, \infty)$ *such that for all* $x \in G$, $b(x) \in \mathcal{C}(\delta)$.

Condition 6.1.5. *There exists* $\underline{\sigma} \in (0, \infty)$ *such that for all* $x \in G$ *and* $\alpha \in \mathbb{R}^m$,

$$\alpha'(\sigma(x)\sigma'(x))\alpha \geq \underline{\sigma}\alpha'\alpha.$$

The following is the main result of [1].

Theorem 6.1.2. *Assume that Conditions 6.1.1-6.1.5 hold. Then the strong Markov process* $\{X^x(\cdot); x \in G\}$ *is positive recurrent and has a unique invariant probability measure.*

We remark that in [1] a somewhat weaker assumption than Condition 6.1.4 is used, which says that $b(x) \in \mathcal{C}(\delta)$ for all x outside a bounded set. In the current work, for simplicity we will use the stronger form as in Condition 6.1.4. Conditions 6.1.1-6.1.5 will be assumed to hold for the rest of this work and will not be explicitly noted in the statements of various results.

We now summarize some of the notation that will be used in this work. For a closed set $G \subset \mathbb{R}^m$, we say $f \in C_b^2(G)$, [respectively $f \in C_c^2(G)$] if f is defined on some open set $O \supset G$ and f is a twice continuously differentiable on O with bounded first two derivatives [respectively compact support]. For $\nu \in \mathcal{P}(S)$ and a ν-integrable $f : S \to \mathbb{R}$, we write $\int_S f d\nu$ as $\langle f, \nu \rangle$ or $\nu(f)$ interchangeably. We will use the symbol "$\Rightarrow$" or "$\xrightarrow{\mathcal{L}}$" to denote convergence in distribution. Let $\mathbb{R}^m$ denote the set of m-dimensional real vectors. Euclidean norm will be denoted by $|\cdot|$ and the corresponding inner product by $\langle \cdot, \cdot \rangle$. The symbols, $\xrightarrow{\mathbb{P}}$, $\xrightarrow{L^p}$ denote convergence in probability and L^p respectively. Denote by $||\cdot||_\infty$ the supremum norm. A vector $v \in \mathbb{R}^m$ is said to be nonnegative (and we write $v \geq 0$) if it is componentwise nonnegative.

6.1.1 Numerical Scheme and Main Results.

Throughout this work, the unique invariant measure for the Markov process $\{X^x\}$ will be denoted by ν. The goal of this work is to develop a convergent numerical procedure for approximating ν. We now describe this procedure.

Let $\{\lambda_k\}_{k\geq 1}$ be a sequence of positive real numbers such that

$$\lambda_k \to 0, \text{ as } k\to\infty \text{ and letting } \Lambda_n := \sum_{k=1}^{n}\lambda_k,\ \Lambda_n\to\infty \text{ as } n\to\infty. \tag{6.1.3}$$

Note the condition is satisfied if $\lambda_n = \frac{1}{n^\theta}$ with $\theta\in(0,1]$. Define the map $\mathcal{S}: G\times\mathbb{R}^m\to G$ by the relation

$$\mathcal{S}(x,v) = \Gamma(x+vi)(1), \tag{6.1.4}$$

where $i:[0,\infty)\to[0,\infty)$ is the identity map. The map $\mathcal{S}$ will be used to construct an Euler discretization of the stochastic dynamical system described by (6.1.2). We now introduce the noise sequence that will be used in the Euler discretization of (6.1.2).

Let $\{U_{k,j}; k\in\mathbb{N}, j=1,...,m\}$ be an array of mutually independent $\mathbb{R}$ valued random variables, given on some probability space $(\Omega,\mathcal{F},\mathbb{P})$, such that $\mathbf{E}U_{k,j}=0$ and $\mathbf{E}U_{k,j}^2=1$, for all $k\in\mathbb{N}, j=1,...,m$. We denote the $\mathbb{R}^m$ valued random variable $(U_{k,1},...,U_{k,m})'$ by U_k. We will make the following assumption on the array $\{U_{k,j}\}$.

Condition 6.1.6. *For some $\alpha\in(0,\infty)$,*

$$\mathbf{E}e^{\lambda U_{k,j}} \leq e^{\alpha\lambda^2} \textit{ for all } k\in\mathbb{N},\ j=1,...,m,\ \lambda\in\mathbb{R}.$$

The above condition is clearly satisfied when $U_{k,j}\sim N(0,1)$. Also, using well known concentration inequalities it can be checked that the condition also holds if $\mathrm{supp}(U_{k,j})$ is uniformly bounded (see Appendix for a proof of the latter statement). Condition 6.1.6 will be assumed to hold throughout this work.

The Euler scheme is given as follows. Define iteratively, sequences $\{X_k\}_{k\in\mathbb{N}_0}$, $\{Y_k\}_{k\in\mathbb{N}_0}$ of G and $\mathbb{R}^m$ valued random variables, respectively, as follows. Fix $x_0\in G$.

$$\begin{cases} X_0 = x_0, \\ Y_{k+1} = X_k + b(X_k)\lambda_{k+1} + \sigma(X_k)\sqrt{\lambda_{k+1}}U_{k+1}, \\ X_{k+1} = \mathcal{S}(X_k, Y_{k+1}-X_k). \end{cases} \tag{6.1.5}$$

Note that $\{X_k\}$ is a sequence of G valued random variables. The last equation of the above display describes a projection for the Euler step that is consistent with the Skorohod problem associated with the problem data.

Define a sequence of $\mathcal{P}(G)$ valued random variables as

$$\nu_n = \frac{1}{\Lambda_n} \sum_{k=1}^{n} \lambda_k \delta_{X_{k-1}}, \quad n \in \mathbb{N}.$$

The above random measures define our basic sequence of approximations for the invariant measure ν. In particular, they yield an approximation for any integral of the form $\int_G f(x) d\nu(x)$ through the corresponding weighted averages:

$$\frac{1}{\Lambda_n} \sum_{k=1}^{n} \lambda_k f(X_{k-1}). \tag{6.1.6}$$

The following is the first main result of this work.

Theorem 6.1.3. *As $n \to \infty$, ν_n converges weakly to ν, almost surely.*

The above result ensures that (6.1.6) gives an almost surely consistent approximation for $\nu(f)$ for any bounded and continuous f. In fact we have a substantially stronger statement as follows:

Theorem 6.1.4. *There exists a $\zeta \in (0, \infty)$ such that for all continuous $f : G \to \mathbb{R}$ satisfying $\limsup_{x\to\infty} e^{-\zeta|x|}|f(x)| = 0$, we have $\nu_n(f) \to \nu(f)$, a.s.*

The key ingredient in the proof of the above almost sure limit theorems is a certain Lyapunov function that was introduced in [19] to study geometric ergodicity properties of reflected diffusions. Using this Lyapunov function we establish a.s. bounds on exponential moments of ν_n that are uniform in n. These bounds in particular guarantee tightness of $\{\nu_n(\omega), n \geq 1\}$, for a.e. ω. Then the remaining work, for proving the above theorems, lies in the characterization of the limit points of $\nu_n(\omega)$. For this we use an extension of the well known Echeverria criterion for invariant distributions of Markov processes that has been developed in [26, 54] (see also [10]). Verification of this criteria (stated as Theorem 6.2.1 in the current work) for a typical limit point ν_0 of $\{\nu_n\}$ requires showing that, ν_0 along with a certain collection $\{\mu_0^i, i = 1, \cdots N\}$ of finite measures supported on various parts of the boundary of G satisfy a relation of the form in (6.2.13). The measures $\{\mu_0^i\}$ are obtained by taking weak limits of certain finite measures constructed from the Euler scheme. Although these pre-limit measures may place positive mass away from the boundary of the domain, we argue using the regularity properties of the Skorohod map (a key ingredient here is Lemma 6.1.1), that in the limit these finite measures are supported on the correct parts of the boundary.

Under additional assumptions, one can obtain rates of convergence as follows. For $\alpha > 0$, set

$$\Lambda_n^{(\alpha)} = \lambda_1^\alpha + ... + \lambda_n^\alpha.$$

Denote the normal distribution with mean a and variance b^2 by $\mathcal{N}(a, b^2)$. For $\phi \in C^3(G)$ (space of three times continuously differentiable functions on G) and $v \in \mathbb{R}^m$, let $D^3\phi(x)(v)^{\otimes 3} = \sum_{i,j,k} D^3_{i,j,k}\phi(x)v_iv_jv_k$.

For $f \in C_c^2(G)$, define $\mathcal{A}f : G \to \mathbb{R}$ and $D_if : G \to \mathbb{R}$; $i = 1, ..., N$ as

$$\mathcal{A}f(x) = b(x) \cdot \nabla f(x) + \frac{1}{2}\sigma'(x)D^2 f(x)\sigma(x), \quad x \in G,$$

$$D_i f(x) = d_i \cdot \nabla f(x), \quad x \in G,$$

where ∇ is the gradient operator and D^2 is the $m \times m$ Hessian matrix.

Theorem 6.1.5. *Assume that U_i's are i.i.d with common distribution μ. There exists a $\zeta \in (0, \infty)$ such that whenever $\phi \in C^2(G)$ satisfies $\lim_{|x|\to\infty} e^{-\zeta|x|}|\nabla\phi(x)|^2 = 0$, we have the following:*

(a) Fast-decreasing step. Suppose $\lim_{n\to\infty} \frac{\Lambda_n^{(3/2)}}{\sqrt{\Lambda_n}} = 0$, $D^2\phi$ is bounded and Lips-

chitz, and

$$\begin{cases} \langle \nabla\phi(x), d_i\rangle = 0, & \forall x \in F_i, \forall i; \\ D^2\phi(x)d_i = \mathbf{0}, & \forall x \in F_i, \forall i. \end{cases} \tag{6.1.7}$$

Then the following CLT holds:

$$\sqrt{\Lambda_n}\nu_n(\mathcal{A}\phi) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_G |\sigma^T\nabla\phi|^2 d\nu\right).$$

(b) Slowly decreasing step. Suppose that $\lim_{n\to\infty}(1/\sqrt{\Lambda_n})\Lambda_n^{(3/2)} = \tilde{\lambda} \in (0, +\infty]$, $\phi \in C^3(G)$ *and* $D^3\phi$ *is bounded and Lipschitz. Further suppose that*

$$\begin{cases} \langle \nabla\phi(x), d_i\rangle = 0, & \forall x \in F_i, \forall i; \\ D^2\phi(x)d_i = \mathbf{0}, & \forall x \in F_i, \forall i; \\ D^3_{\cdot jk}\phi(x)\cdot d_i = 0, & \forall x \in F_i, \forall i, j, k. \end{cases} \tag{6.1.8}$$

Then we have

$$\sqrt{\Lambda_n}\nu_n(\mathcal{A}\phi) \xrightarrow{\mathcal{L}} \mathcal{N}\left(\tilde{\lambda}\tilde{m}, \int_G |\sigma^T\nabla\phi|^2 d\nu\right) \qquad \text{if } \tilde{\lambda} < \infty, \tag{6.1.9}$$

$$\frac{\Lambda_n}{\Lambda_n^{(3/2)}}\nu_n(\mathcal{A}\phi) \xrightarrow{\mathbb{P}} \tilde{m} \qquad \text{if } \tilde{\lambda} = +\infty, \tag{6.1.10}$$

where

$$\tilde{m} = -\frac{1}{6}\int_G\int_{\mathbb{R}^m} D^3\phi(x)(\sigma(x)u)^{\otimes 3}\mu(du)\nu(dx).$$

Note that when $\lambda_k = \frac{1}{k^\alpha}$, $\Lambda_n^{(3/2)}/\sqrt{\Lambda_n}$ converges to 0 [resp. ∞, $\tilde{\lambda} \in (0, +\infty)$], if $\alpha > 1/2$ [resp. $\alpha < 1/2$, $\alpha = 1/2$]. Also note that if ϕ is a smooth function supported in the interior of G then it automatically satisfies (6.1.7) and (6.1.8). Our assumptions are restrictive in the sense that it does not cover the functional $f(x) = x_1$, but note that the theorem does cover a function $f_\epsilon(x)$ which is given as a suitable mollification of the function $x_1\prod_{i=1}^m 1_{\{x_i \geq \epsilon\}}$ that vanishes smoothly at the boundaries. Heuristically speaking, this functional captures some properties of the functional $f(x) = x_1$ when ϵ is small. Another functional of interest for which the theorem gives a rate of convergence is a suitable mollification of the functional $\prod_{i=1}^{m} 1_{\{x_i \geq \alpha_i\}}$ where $\alpha_i > 0$, $i = 1, \cdots m$.

Proof of Theorem 6.1.5 is quite similar to that of Theorem 9 in [57], the main difference is in the treatment of the reflection terms for which once more we appeal to regularity properties of the Skorohod map and an estimate based on Lemma 6.1.1 (see proof of (6.2.19) which is crucially used in proofs of Section 6.3).

A key step in the implementation of the Euler scheme in (6.1.5) is the evaluation of the one time step Skorohod map $\mathcal{S}(x, v)$. In Section 6.4.1 we describe one possible approach to this evaluation that uses relationships between Skorohod problems and Linear Complementarity problems(LCPs). There are many well developed numerical codes for solving LCPs (for example in MATLAB) and we will describe in Section 6.4.2 some results from numerical experiments that use a quadratic programming algorithm for LCPs (cf. [24]) in implementing the scheme in (6.1.5). As remarked earlier, one of the advantages of Monte-Carlo methods is the ease of implementation, particularly for high dimensional problems. To illustrate this, in Section 6.4.2 we present numerical results for a eight dimensional Skorohod problem. Comparisons with known exact formulas (from [25]) for this problem show that the scheme performs well for small correlation values.

This chapter is organized as follows. In Section 6.2 we prove Theorem 6.1.3 and 6.1.4. Theorem 6.1.3 is proved in two steps. Section 6.2.1 shows the tightness of the random measures $\{\nu_n\}$, and Section 6.2.2 characterizes the limit of the measures $\{\nu_n\}$ as the invariant measure of the constrained diffusion in (6.1.1). Section 6.2.3 gives the proof of Theorem 6.1.4. Rate of convergence theorem (Theorem 6.1.5) is proved in Section 6.3. Finally we conclude by describing some numerical results in Section 6.4.

6.2 Proofs of Theorems 6.1.3 and 6.1.4.

The proof of Theorem 6.1.3 proceeds by showing that for a.e. ω, the sequence of random probability measures $\{\nu_n(\omega)\}_{n\geq 1}$ is tight and then characterizing the limit points of the sequence using a generalization of Echeverria's criteria. Tightness is argued in Section 6.2.1 while the limit points are characterized in Section 6.2.2. Finally in Section 6.2.3, we give the proof of Theorem 6.1.4.

6.2.1 Tightness.

We begin by presenting a Lyapunov function introduced in [1] that plays a key role in the stability analysis of constrained diffusion processes of the form studied here (see [1, 19, 20, 21, 10, 8, 7]).

Throughout this work we will fix a $\delta > 0$ as in Condition 6.1.4.

For $x \in G$, let $\mathcal{A}(x)$ be the collection of all absolutely continuous functions $z : [0,\infty) \to \mathbb{R}^m$ defined via

$$z(t) \doteq \Gamma\left(x + \int_0^{\cdot} v(s)ds\right)(t), \quad t \in [0,\infty), \tag{6.2.1}$$

for some $v : [0,\infty) \to \mathcal{C}(\delta)$ which satisfies

$$\int_0^t |v(s)|ds < \infty, \quad \text{for all } t \in [0,\infty). \tag{6.2.2}$$

Namely,

$$\mathcal{A}(x) = \left\{ z : [0,\infty) \to \mathbb{R}^m \;\middle|\; \begin{array}{l} z \text{ is absolutely continuous,} \\ (6.2.1) \text{ and } (6.2.2) \text{ hold for some } v : [0,\infty) \to \mathcal{C}(\delta) \end{array} \right\}.$$

Define $T : G \to [0,\infty)$ by the relation

$$T(x) \doteq \sup_{z\in\mathcal{A}(x)} \inf\{t \in [0,\infty) : z(t) = 0\}, \quad x \in G. \tag{6.2.3}$$

The function T has the following properties (see [1]).

Lemma 6.2.1. *There exist constants $c, C \in (0, \infty)$ such that the following hold:*

(i) *For all $x, y \in G$,*
$$|T(x) - T(y)| \leq C|x - y|.$$

(ii) *For all $x \in G$, $T(x) \geq c|x|$. Thus, in particular, for all $M \in (0, \infty)$ the set $\{x \in G : T(x) \leq M\}$ is compact.*

(iii) *Fix $x \in G$ and let $z \in \mathcal{A}(x)$. Then for all $t > 0$,*
$$T(z(t)) \leq (T(x) - t)^+.$$

We next present an elementary lemma that will be used in obtaining moment estimates. For $k \in \mathbb{N}$, let $\mathcal{F}_k = \sigma(U_1, ..., U_k)$. Set $\mathcal{F}_0 = \{\emptyset, \Omega\}$.

Lemma 6.2.2. *There exist $c_1, c_2 \in (1, \infty)$ for which the following holds. Let $\{v_i\}_{i \in \mathbb{N}}$ be a sequence of $\mathbb{R}^m$ valued random variables such that v_i is $\mathcal{F}_{i-1}$ measurable for all $i \geq 1$ and*
$$\operatorname*{ess\,sup}_{\omega} |v_i(\omega)| \equiv |v_i|_\infty < \infty.$$
Let $S_n = \sum_{i=1}^n v_i \cdot U_i$, $n \in \mathbb{N}$. Then for every $r \geq 0$ and $n \geq 1$,
$$\mathbf{E} \max_{1 \leq i \leq n} e^{r|S_i|} \leq c_1 e^{c_2 r^2 \sum_{i=1}^n |v_i|_\infty^2}.$$

Proof. We will only give the proof for the case $m = 1$. The general case is treated similarly.

From Doob's maximal inequalities for submartingales, we have
$$\begin{aligned}\mathbf{E} \max_{1 \leq i \leq n} e^{r|S_i|} &\leq 4\mathbf{E} e^{r|S_n|} \\ &\leq 4\left(\mathbf{E} e^{rS_n} + \mathbf{E} e^{-rS_n}\right)\end{aligned}$$

From Condition 6.1.6, it follows that for every $r \in \mathbb{R}$,

$$\begin{aligned}\mathbf{E}\left(e^{rS_n}|\mathcal{F}_{n-1}\right) &\leq e^{rS_{n-1}}e^{\alpha r^2 v_n^2} \\ &\leq e^{rS_{n-1}}e^{\alpha r^2 |v_n|_\infty^2}\end{aligned}$$

The result now follows by a successive conditioning argument. □

Define $\lambda : [0,\infty) \to [0,\infty)$ and $j : [0,\infty) \to \mathbb{N}_0$ as

$$\lambda(s) = \Lambda_k; j(s) = k, \text{ if } \Lambda_k \leq s < \Lambda_{k+1}, \ k \in \mathbb{N}_0;$$

where we define $\Lambda_0 = 0$. Define piecewise linear $\mathbb{R}^m$ valued stochastic process as follows,

$$\hat{W}(t) = \sum_{i \leq j(t)} \sqrt{\lambda_i} U_i + \frac{t - \lambda(t)}{\sqrt{\lambda_{j(t)+1}}} U_{j(t)+1}, \quad t \geq 0.$$

Let $\hat{X}(t)$ be the solution of the following integral equation

$$\hat{X}(t) = \Gamma\left(x_0 + \int_0^{\cdot} b(\hat{X}(\lambda(s)))ds + \int_0^{\cdot} \sigma(\hat{X}(\lambda(s)))d\hat{W}(s)\right)(t), \quad t \geq 0.$$

Clearly, $\hat{X}(\lambda(t)) = X_{j(t)}$ for all $t \geq 0$.

Fix $\rho \in (0,1]$. Define

$$\varpi = \frac{1}{2(1+\rho)L}, \quad \Delta = 4\lambda_0 + 16L\ln(c_1), \tag{6.2.4}$$

where $L = c_2 a_1^2 C^2 K^2$ and $\lambda_0 = \sup_{i\geq 1} \lambda_i$. Let $V : G \to \mathbb{R}_+$ be defined as

$$V(x) = e^{\varpi T(x)}, \quad x \in G.$$

Lemma 6.2.3. *There exist $\beta \in (0,1)$ and $\kappa \in [0,\infty)$ such that for each $\zeta \in [0,\rho]$ and for all $t \geq 0$,*

$$\mathbf{E}(V^{1+\zeta}(X_{j(t+\Delta)})|\mathcal{F}_{j(t)}) \leq (1-\beta)V^{1+\zeta}(X_{j(t)}) + \kappa \tag{6.2.5}$$

Proof. Fix $t \geq 0$ and $\zeta \in [0, \rho]$. Define $\xi : [\lambda(t), \infty) \to G$ as

$$\xi(s) = \Gamma\left(X_{j(t)} + \int_{\lambda(t)}^{\lambda(t)+\cdot} b(X_{j(u)})du\right)(s - \lambda(t)), \quad s \geq \lambda(t).$$

Using the Lipschitz property of the Skorokhod map (Condition 6.1.1), we have

$$\begin{aligned}\sup_{\lambda(t) \leq s \leq \lambda(t)+\Delta+\lambda_0} |\hat{X}(s) - \xi(s)| \leq & K \sup_{\lambda(t) \leq s \leq \lambda(t)+\Delta+\lambda_0} \left| \int_{\lambda(t)}^{s} \sigma(\hat{X}(\lambda(u)))d\hat{W}(u) \right| \\ =: & K\bar{\nu}(t, \Delta).\end{aligned}$$

Note that

$$\Delta - \lambda_0 \leq \lambda(t + \Delta) - \lambda(t) \leq \Delta + \lambda_0.$$

Using this observation along with Lemma 6.2.1 (i) and (iii),

$$\begin{aligned}T(\hat{X}(\lambda(t+\Delta))) \leq & T(\xi(\lambda(t+\Delta))) + CK\bar{\nu}(t,\Delta) \\ \leq & (T(\hat{X}(\lambda(t))) - (\lambda(t+\Delta) - \lambda(t)))^+ + CK\bar{\nu}(t,\Delta) \\ \leq & (T(\hat{X}(\lambda(t))) - (\Delta - \lambda_0))^+ + CK\bar{\nu}(t,\Delta).\end{aligned}$$

From the above estimate and the definition of $V(x)$, we now have

$$\begin{aligned}&\frac{\mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))^{1+\zeta}|\mathcal{F}_{j(t)})}{V(\hat{X}(\lambda(t)))^{1+\zeta}} \\ &\leq \mathbf{E}\left(\exp(\varpi(1+\zeta)\left((T(\hat{X}(\lambda(t))) - (\Delta - \lambda_0))^+ + CK\bar{\nu}(t,\Delta)\right)\right) |\mathcal{F}_{j(t)}\Big) \\ &\quad \times \exp(-\varpi(1+\zeta))T(\hat{X}(\lambda(t))). \qquad (6.2.6)\end{aligned}$$

Letting, for $q \in \mathbb{N}_0$, $\sigma_q = \sigma(X_q)$, we have, for any $s \in [\lambda(t), \lambda(t) + \Delta + \lambda_0]$,

$$\int_{\lambda(t)}^{s} \sigma(\hat{X}(\lambda(u)))d\hat{W}(u) \leq \begin{cases} \sum_{q=j(t)}^{j(s)} \sigma_q \sqrt{\lambda_{q+1}} U_{q+1}, & \text{if } \sigma_{j(s)} U_{j(s)} \geq 0 \\ \\ \sum_{q=j(t)}^{j(s)-1} \sigma_q \sqrt{\lambda_{q+1}} U_{q+1}, & \text{if } \sigma_{j(s)} U_{j(s)} < 0 \end{cases},$$

which can be bounded by

$$\max_{j(t) \leq j \leq j(s)} \sum_{q=j(t)}^{j} \sigma_q \sqrt{\lambda_{q+1}} U_{q+1}.$$

Similarly,

$$-\int_{\lambda(t)}^{s}\sigma(\hat{X}(\lambda(u)))d\hat{W}(u) \leq \max_{j(t)\leq j\leq j(s)} -\sum_{q=j(t)}^{j}\sigma_q\sqrt{\lambda_{q+1}}U_{q+1}.$$

And therefore

$$\bar{\nu}(t,\Delta) = \sup_{\lambda(t)\leq s\leq \lambda(t)+\Delta+\lambda_0}\left|\int_{\lambda(t)}^{s}\sigma(\hat{X}(\lambda(u)))d\hat{W}(u)\right| \leq \max_{j(t)\leq j\leq j_t^*}\left|\sum_{q=j(t)}^{j}\sigma_q\sqrt{\lambda_{q+1}}U_{q+1}\right|,$$

where $j_t^* = j(\lambda(t)+\Delta+\lambda_0)$.

Using Lemma 6.2.2, we now have that, with $m_0 = \varpi(1+\zeta)CK$,

$$\mathbf{E}\left[e^{m_0\bar{\nu}(t,\Delta)}\,\middle|\mathcal{F}_{j(t)}\right] \leq c_1 e^{c_2 m_0^2 a_1^2 \sum_{q=j(t)}^{j_t^*}\lambda_{q+1}} \leq c_1 e^{c_2 m_0^2 a_1^2(\Delta+2\lambda_0)}. \tag{6.2.7}$$

In the case $T(\hat{X}(\lambda(t))) \geq \Delta - \lambda_0$, we have from (6.2.6) and (6.2.7) that

$$\mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))^{1+\zeta}|\mathcal{F}_{j(t)}) \leq V(\hat{X}(\lambda(t)))^{1+\zeta}e^{-\varpi(1+\zeta)(\Delta-\lambda_0)} \times c_1 e^{c_2 m_0^2 a_1^2(\Delta+2\lambda_0)}.$$

Recalling the choice of ϖ and Δ, we now see that

$$\mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))^{1+\zeta}|\mathcal{F}_{j(t)}) \leq (1-\beta)V(\hat{X}(\lambda(t)))^{1+\zeta},$$

where $\beta = 1 - e^{-3\ln c_1}$.

In the case $T(\hat{X}(\lambda(t))) < \Delta - \lambda_0$, we have

$$\begin{aligned}\mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))^{1+\zeta}|\mathcal{F}_{j(t)}) &\leq \mathbf{E}\left(e^{m_0\bar{\nu}(t,\Delta)}|\mathcal{F}_{j(t)}\right)\\ &\leq c_1 e^{c_2 m_0^2 a_1^2(\Delta+2\lambda_0)} \leq c_1 e^{\frac{1}{4L}(\Delta+2\lambda_0)} \equiv \kappa.\end{aligned}$$

Combining the two cases, we have (6.2.5). $\square$

The following lemma follows from Lemma 6.2.3 through a recursive argument.

Lemma 6.2.4. *There exists $a_2 \in (0,\infty)$ such that*

$$\sup_t \mathbf{E}(V(\hat{X}(\lambda(t)))^{1+\rho}) \leq a_2. \tag{6.2.8}$$

Proof. For any $t \in (\Delta, \infty)$, we can find $t' \in (0, \Delta]$ and $j \in \mathbb{N}$ such that $t = t' + j\Delta$. By a recursive argument using (6.2.5), we then have

$$\mathbf{E}(V(\hat{X}(\lambda(t)))^{1+\rho}) \leq \mathbf{E}(V(\hat{X}(\lambda(t')))^{1+\rho}) + \frac{\kappa}{\beta}.$$

Thus

$$\sup_t \mathbf{E}(V(\hat{X}(\lambda(t)))^{1+\rho}) \leq \sup_{0 \leq t < \Delta} \mathbf{E}(V(\hat{X}(\lambda(t)))^{1+\rho}) + \frac{\kappa}{\beta}.$$

The supremum on the right side is bounded by $\max_{j \leq j(\Delta+\lambda_0)} \mathbf{E}(V(X_j)^{1+\rho})$, which is finite using Condition 6.1.6, boundedness of b, σ and the Lipschitz property of Γ. $\square$

Now we can prove the following lemma.

Lemma 6.2.5. *For a.e.* ω, $\sup_n \langle V, \nu_n(\omega) \rangle < \infty$. *Consequently, the sequence* $\{\nu_n(\omega)\}_{n \geq 1}$ *is tight for a.e.* ω.

Proof. Let n_0 be such that $\Lambda_{n_0} > \Delta$. Then it suffices to consider the supremum in the above display over all $n \geq n_0$. For $i \in \mathbb{N}_0$, define $s(i) = \inf\{j \in \mathbb{N}_0 : \Lambda_j \geq i\Delta\}$. Then $s(\lfloor \Lambda_n/\Delta \rfloor) \leq n$ and therefore, for $n \geq n_0$,

$$\begin{aligned}\nu_n(V) =& \frac{1}{\Lambda_n} \sum_{k=1}^{n} \lambda_k V(X_{k-1}) = \frac{1}{\Lambda_n} \int_0^{\Lambda_n} V(\hat{X}(\lambda(t)))dt \\ \leq& \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor + 1)\Delta} V(\hat{X}(\lambda(t)))dt.\end{aligned}$$

Using Lemma 6.2.3 with $\zeta = 0$, we have

$$\begin{aligned}&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor + 1)\Delta} V(\hat{X}(\lambda(t)))dt \\ &\quad \leq \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor + 1)\Delta} [V(\hat{X}(\lambda(t))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt \\ &\quad + \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor + 1)\Delta} [V(\hat{X}(\lambda(t)))(1-\beta) + \kappa]dt.\end{aligned}$$

Thus, rearranging terms,

$$\begin{aligned}
&\frac{\beta}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} V(\hat{X}(\lambda(t)))dt \\
&\qquad \le \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt \\
&\qquad + \kappa \frac{(\lfloor \Lambda_n/\Delta \rfloor + 1)\Delta}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}}.
\end{aligned}$$

Next note that $\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)} \ge \lfloor \Lambda_n/\Delta \rfloor \Delta$, and, for $n \ge n_0$,

$$\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)} \ge \Lambda_{s(1)} \ge \lambda_1.$$

Thus

$$\sup_{n \ge n_0} \kappa \frac{(\lfloor \Lambda_n/\Delta \rfloor + 1)\Delta}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \le \kappa(1 + \frac{\Delta}{\lambda_1}) < \infty.$$

To prove the lemma, it is now enough to show

$$\sup_{n \ge n_0} \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt < \infty, \quad \text{a.e. } \omega. \tag{6.2.9}$$

The above expression can be split into two terms:

$$\begin{aligned}
&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt \\
&= \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t))) - V(\hat{X}(\lambda(t+\Delta)))]dt \\
&\quad + \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt \\
&\equiv T_1 + T_2
\end{aligned}$$

Consider the first term:

$$
\begin{aligned}
T_1 =&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t))) - V(\hat{X}(\lambda(t+\Delta)))]dt \\
=&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} V(\hat{X}(\lambda(t)))dt - \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_\Delta^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta+\Delta} V(\hat{X}(\lambda(t)))dt \\
=&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{\Delta} V(\hat{X}(\lambda(t)))dt - \frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta}^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta+\Delta} V(\hat{X}(\lambda(t)))dt \\
\leq&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \sum_{k:\Lambda_{k-1}\leq\Delta} \lambda_k V(X_{k-1}).
\end{aligned}
$$

Let $Z = \sum_{k:\Lambda_{k-1}\leq\Delta} \lambda_k V(X_{k-1})$. Then from (6.2.8), we have $\mathbf{E}Z \leq a_2(\Delta+\lambda_0)$. Combining this with the fact that for $n \geq n_0$, $\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)} \geq \lambda_1$, we have that

$$
\sup_{n\geq n_0} T_1(\omega) < \infty, \quad \text{a.e. } \omega. \tag{6.2.10}
$$

Next, consider T_2:

$$
\begin{aligned}
T_2 =&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \int_0^{(\lfloor \Lambda_n/\Delta \rfloor+1)\Delta} [V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt \\
=&\frac{1}{\Lambda_{s(\lfloor \Lambda_n/\Delta \rfloor)}} \sum_{i=0}^{\lfloor \Lambda_n/\Delta \rfloor} \int_{i\Delta}^{(i+1)\Delta} [V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt.
\end{aligned}
$$

From Kronecker's Lemma (see page 63 of [38]), the last sum is bounded in n a.s. (in fact converges to 0) if the following series is summable a.s.

$$
\sum_{i=1}^{\infty} \frac{1}{\Lambda_{s(i)}} \int_{i\Delta}^{(i+1)\Delta} [V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt.
$$

Consider the sum over even and odd terms separately. For even terms, the sum can be written as

$$
\sum_{k=1}^{\infty} \frac{1}{\Lambda_{s(2k)}} \int_{2k\Delta}^{(2k+1)\Delta} [V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt. \tag{6.2.11}
$$

Let

$$
\xi_{k+1} = \frac{1}{\Lambda_{s(2k)}} \int_{2k\Delta}^{(2k+1)\Delta} [V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt
$$

and $\mathcal{G}_k = \mathcal{F}_{j(2k\Delta)}$, then we have $\mathbf{E}(\xi_{i+1}|\mathcal{G}_i) = 0$. Also note that ξ_{i+1} is $\mathcal{G}_{i+1}$ measurable. Thus $S_n = \sum_{i=1}^n \xi_i$ is a martingale with respect to the filtration $\{\mathcal{G}_n\}$. Consequently, by Chow's Theorem (see Theorem 2.17 of [45]), the series in (6.2.11) is a.s. summable if $\sum_{k=1}^\infty \mathbf{E}(|\xi_k|^{1+\rho}) < \infty$. Now note that

$$\begin{aligned}\mathbf{E}|\xi_k|^{1+\rho} &= \mathbf{E}\left(\left|\frac{1}{\Lambda_{s(2k)}}\int_{2k\Delta}^{(2k+1)\Delta}[V(\hat{X}(\lambda(t+\Delta))) - \mathbf{E}(V(\hat{X}(\lambda(t+\Delta)))|\mathcal{F}_{j(t)})]dt\right|^{1+\rho}\right) \\ &\leq \frac{2^{1+\rho}\Delta^{1+\rho}}{\Lambda_{s(2k)}^{1+\rho}} \sup_t \mathbf{E}(V(\hat{X}(t))^{1+\rho}) \leq \frac{2^{1+\rho}\Delta^{1+\rho}a_2}{\Lambda_{s(2k)}^{1+\rho}},\end{aligned}$$

where the last inequality follows form Lemma 6.2.4. Since $\Lambda_{s(k)} \geq k\Delta$, we have that

$$\sum_{k=1}^\infty \frac{1}{\Lambda_{s(k)}^{1+\rho}} \leq \frac{1}{\Delta^{1+\rho}}\sum_{k=1}^\infty \frac{1}{k^{1+\rho}} < \infty.$$

This proves that the series in (6.2.11) is summable. The odd terms are treated in a similar manner. Thus we have proved

$$\sup_{n\geq n_0} T_2(\omega) < \infty, \quad \text{a.e. } \omega. \tag{6.2.12}$$

Now (6.2.9) is an immediate consequence of (6.2.10) and (6.2.12), which proves the lemma. □

6.2.2 Identification of the limit.

In this section we will complete the proof of Theorem 6.1.3 by arguing that for a.e. ω, every weak limit point of $\nu_n(\omega)$ equals ν. For this we will use the following extension of the Echeverria Criteria (see [54, 82], see also Theorem 5.7 of [10]).

Theorem 6.2.1. *Let $\nu_0 \in \mathcal{P}(G)$ and $\mu_0^i \in \mathcal{M}_F(F_i)$, $i = 1, ..., N$ be such that for all $f \in C_c^2(G)$,*

$$\nu_0(\mathcal{A}f) + \sum_{i=1}^N \mu_0^i(D_i f) = 0. \tag{6.2.13}$$

Then $\nu_0 = \nu$.

In order to apply the above theorem to show convergence of ν_n to ν, we will consider a sequence of finite measure $\{\mu_n^i\}_{n\in\mathbb{N}}$; $i = 1, ..., N$, which, roughly speaking, correspond to the prelimit versions of the measures $\{\mu_0^i\}$ that appear in the theorem above. We now describe this sequence.

For $u \in \mathbb{R}^m,\ v \in G,\ r \in (0, \infty)$, define, for $t \in [0, 1]$,

$$
\begin{aligned}
\mathbf{z}(u, v, r|t) &\equiv \mathbf{z}(t) = v + (b(v)r + \sigma(v)\sqrt{r}u)t,\\
\mathbf{x}(u, v, r|t) &\equiv \mathbf{x}(t) = \Gamma(\mathbf{z})(t),\\
\mathbf{y}(u, v, r|t) &\equiv \mathbf{y}(t) = \mathbf{x}(t) - \mathbf{z}(t).
\end{aligned}
$$

Then, one can represent the trajectory $\mathbf{y}$ as

$$
\mathbf{y}(t) = \sum_{i=1}^{N} d_i \int_0^t \alpha_i(s) d|\mathbf{y}|(s); \quad t \in [0, 1], \tag{6.2.14}
$$

where $\alpha_i(s) \equiv \alpha_i(u, v, r|s) \in [0, 1]$ and $\alpha_i(s) > 0$ only if $\mathbf{x}(s) \in F_i$. Also, let, for $t \in [0, 1]$

$$
\begin{aligned}
\mathbf{\Pi^t}(u, v, r) &= \mathbf{z}(1) + t(\mathbf{x}(1) - \mathbf{z}(1)),\\
\mathbf{L^i}(u, v, r) &= \int_0^1 \alpha_i(t) d|\mathbf{y}|(t), \quad i = 1, ..., N.
\end{aligned}
$$

Finally for $k \in \mathbb{N}_0$, let

$$
\Pi_k^t = \mathbf{\Pi^t}(U_{k+1}, X_k, \lambda_{k+1}),\ L_k^i = \mathbf{L^i}(U_{k+1}, X_k, \lambda_{k+1}).
$$

For $k \in \mathbb{N}_0$ and $i = 1, ..., N$, define a $\mathcal{M}_F(\mathbb{R}^m)$ valued random variable m_k^i by the relation

$$
\langle \psi, m_k^i \rangle = \int_0^1 \mathbf{E}_{X_k}[\psi(\Pi_k^t) L_k^i] dt, \quad \psi \in BM_+(\mathbb{R}^m), \tag{6.2.15}
$$

where $\mathbf{E}_X[Z]$ denotes $\mathbf{E}[Z|X]$, and $BM_+(\mathbb{R}^m)$ is the space of nonnegative bounded measurable functions on $\mathbb{R}^m$.

For $n \in \mathbb{N}$ and $i = 1, ..., N$, let μ_n^i be a $\mathcal{M}_F(\mathbb{R}^m)$ valued random variable defined as

$$\mu_n^i(A) = \frac{1}{\Lambda_n} \sum_{k=0}^{n-1} m_k^i(A); \quad A \in \mathcal{B}(\mathbb{R}^m).$$

The following lemma relates the above family of random measures with our approximation scheme. Recall the definition of the filtration $\{\mathcal{F}_k\}$ in Section 6.2.1.

Lemma 6.2.6. *For every $f \in C_b^2(\mathbb{R}^m)$, there exists a sequence of real random variables $\{\xi_n^f\}_{n \in \mathbb{N}}$ such that*

$$\frac{1}{\Lambda_n} \sum_{k=0}^{n-1} \mathbf{E}[f(X_{k+1}) - f(X_k) | \mathcal{F}_k] = \sum_{i=1}^{N} \mu_n^i(D_i f) + \nu_n(\mathcal{A} f) + \xi_n^f, \tag{6.2.16}$$

and $\sup_n \xi_n^f(\omega) < \infty$ a.s. Furthermore if f has compact support then $\xi_n^f \to 0$ a.s. as $n \to \infty$.

Proof. Fix $(u, v, r) \in \mathbb{R}^m \times G \times (0, \infty)$. Using the notation introduced above, we have from Taylor's theorem,

$$f(z(1)) - f(v) = \langle \nabla f(v), \eta \rangle + \frac{1}{2} \eta' D^2 f(v) \eta + R_2(v, z(1))$$

where

$$R_2(x, y) = f(y) - f(x) - \langle \nabla f(x), y - x \rangle - \frac{1}{2}(y - x)^T D^2 f(x)(y - x)$$

and

$$\eta \equiv \eta(u, v, r) = b(v) r + \sigma(v) \sqrt{r} u.$$

Define

$$r_2(x, y) = \frac{1}{2} \sup_{t \in (0,1)} ||D^2 f(x + t(y - x)) - D^2 f(x)||,$$

then we have $|R_2(x, y)| \leq r_2(x, y) |x - y|^2$.

Also

$$\begin{aligned}
f(\mathbf{x}(1)) - f(\mathbf{z}(1)) &= \int_0^1 \frac{df(\mathbf{z}(1) + t(\mathbf{x}(1) - \mathbf{z}(1)))}{dt} dt \\
&= \int_0^1 \nabla f(\mathbf{z}(1) + t(\mathbf{x}(1) - \mathbf{z}(1))) \cdot (\mathbf{x}(1) - \mathbf{z}(1)) dt \\
&= \sum_{i=1}^{N} \int_0^1 \nabla f(\mathbf{z}(1) + t(\mathbf{x}(1) - \mathbf{z}(1))) dt \cdot d_i \int_0^1 \alpha_i(t) d|\mathbf{y}|(t).
\end{aligned}$$

Fix a $k \in \mathbb{N}$ and let $v = X_k$, $u = U_{k+1}$ and $r = \lambda_{k+1}$. Then

$$\mathbf{E}[f(X_{k+1}) - f(X_k)|\mathcal{F}_k] = \mathbf{E}[f(\mathbf{x}(1)) - f(\mathbf{z}(1)) + f(\mathbf{z}(1)) - f(v)|\mathcal{F}_k].$$

From the definition of m_k^i in (6.2.15) and observing that $\{X_k, \mathcal{F}_k\}$ is a Markov chain and U_{k+1} is independent of $\mathcal{F}_k$, it follows that

$$\mathbf{E}[f(\mathbf{x}(1)) - f(\mathbf{z}(1))|\mathcal{F}_k] = \sum_{i=1}^{N} m_k^i(D_i f),$$

Using independence of U_{k+1} from $\mathcal{F}_k$ once more,

$$\begin{aligned}
\mathbf{E}[f(\mathbf{z}(1)) - f(v)|\mathcal{F}_k] =& \lambda_{k+1}\langle \nabla f(X_k), b(X_k)\rangle + \frac{1}{2}\lambda_{k+1}\sigma(X_k)' D^2 f(X_k)\sigma(X_k) \\
&+ \frac{1}{2}\lambda_{k+1}^2 b(X_k)' D^2 f(X_k) b(X_k) + \mathbf{E}[R_2(X_k, X_k + \eta_k)|\mathcal{F}_k] \\
=& \lambda_{k+1}\mathcal{A}f(X_k) + \xi^f(k),
\end{aligned}$$

where

$$\xi^f(k) = \frac{1}{2}\lambda_{k+1}^2 b(X_k)' D^2 f(X_k) b(X_k) + \mathbf{E}[R_2(X_k, X_k + \eta_k)|\mathcal{F}_k]$$

and $\eta_k = \eta(U_{k+1}, X_k, \lambda_{k+1})$.

Thus we have

$$\begin{aligned}
&\frac{1}{\Lambda_n}\sum_{k=0}^{n-1} \mathbf{E}[f(X_{k+1}) - f(X_k)|\mathcal{F}_k] \\
&\qquad = \frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\left[\sum_{i=1}^{N} m_k^i(D_i f) + \lambda_{k+1}\mathcal{A}f(X_k) + \xi^f(k)\right] \\
&\qquad = \sum_{i=1}^{N} \mu_n^i(D_i f) + \nu_n(\mathcal{A}f) + \frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\xi^f(k).
\end{aligned}$$

Equality in (6.2.16) follows on taking $\xi_n^f = \frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\xi^f(k)$.

We now show that $\sup_n \xi_n^f(\omega) < \infty$ a.s. Write

$$\xi^f(k) = \frac{1}{2}\lambda_{k+1}^2 b(X_k)' D^2 f(X_k) b(X_k) + \mathbf{E}[R_2(X_k, X_k+\eta_k)|\mathcal{F}_k] \equiv \xi_1^f(k) + \xi_2^f(k).$$

The term $\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\xi_1^f(k)$ converges to zero because of the boundedness of b and $D^2 f$. Consider now the contribution from $\xi_2^f(k)$. Let for $p \in \mathbb{R}_+$,

$$h(p) = \frac{1}{2}\sup_{\substack{x_1,x_2\in\mathbb{R}^m\\ |x_1-x_2|\leq p}} ||D^2 f(x_2) - D^2 f(x_1)||.$$

Then

$$|R_2(X_k, X_k+\eta_k)| \leq h(\eta_k)|\eta_k|^2$$

and so for some $\kappa_1 \in (0,\infty)$,

$$\begin{aligned}|\xi_2^f(k)| &\leq \mathbf{E}[h(\eta_k)|\eta_k|^2|\mathcal{F}_k]\\ &\leq ||h||_\infty \kappa_1 \lambda_{k+1}.\end{aligned}$$

Thus $\sup_n \xi_n^f(\omega) < \infty$ a.s. This completes the first part of the lemma.

Finally if f in addition has compact support, we have $h(p) \to 0$ as $p \to 0$. Fix $\epsilon > 0$. Since b, σ are bounded, we can find for each $\theta \in (0,\infty)$, $k_\theta \in \mathbb{N}$ such that for every $k \geq k_\theta$,

$$|h(b(x_k)\lambda_{k+1} + \sigma(x_k)\sqrt{\lambda_{k+1}}U_{k+1})|1_{|U_{k+1}|\leq\theta} \leq \epsilon.$$

Also, for some $l_\eta \in (0,\infty)$, for all $k \in \mathbb{N}$,

$$\mathbf{E}[|\eta_k|^2 1_{|U_{k+1}|\geq\theta}|\mathcal{F}_k] \leq l_\eta(\lambda_k^{3/2} + \lambda_k \mathbf{E}[|U_1|^2 1_{|U_1|\geq\theta}]) \quad \text{a.s.},$$

$$\mathbf{E}[|\eta_k|^2|\mathcal{F}_k] \leq l_\eta \lambda_k \quad \text{a.s.}$$

Choose $\theta_0 \in (0,\infty)$ such that $\mathbf{E}[|U_1|^2 1_{|U_1|\geq\theta_0}] \leq \epsilon$. Then

$$\frac{1}{\Lambda_n}\sum_{k=k_{\theta_0}}^{n-1}|\xi_2^f(k)| \leq \epsilon l_\eta \frac{1}{\Lambda_n}\sum_{k=k_{\theta_0}}^{n-1}\lambda_k + ||h||_\infty l_\eta(\frac{1}{\Lambda_n}\sum_{k=k_{\theta_0}}^{n-1}\lambda_k^{3/2} + \frac{\epsilon}{\Lambda_n}\sum_{k=k_{\theta_0}}^{n-1}\lambda_k).$$

Thus,

$$\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}|\xi_2^f(k)| \leq \frac{1}{\Lambda_n}\sum_{k=0}^{k_{\theta_0}-1}\xi_2^f(k) + \epsilon l_\eta(1+||h||_\infty) + ||h||_\infty l_\eta \frac{1}{\Lambda_n}\sum_{k=k_{\theta_0}}^{n-1}\lambda_k^{3/2}.$$

Sending $n \to \infty$ and then $\epsilon \to 0$, we now see that $\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}|\xi_2^f(k)| \to 0$ as $n \to \infty$. The result follows. □

The following lemma shows that the left side of the expression in (6.2.16) converges to 0 as $n \to \infty$.

Lemma 6.2.7. *For every $f \in C_b^2(G)$,*

$$\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\mathbf{E}[f(X_{k+1}) - f(X_k)|\mathcal{F}_k] \to 0 \ a.s., \quad as\ n \to \infty.$$

Proof. We can split the sum into two terms:

$$\begin{aligned}
&\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\mathbf{E}[f(X_{k+1}) - f(X_k)|\mathcal{F}_k] \\
&\quad =\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}(\mathbf{E}[f(X_{k+1})|\mathcal{F}_k] - f(X_{k+1})) + \frac{1}{\Lambda_n}\sum_{k=0}^{n-1}(f(X_{k+1}) - f(X_k)) \\
&\quad =T_1 + T_2.
\end{aligned}$$

Note that,

$$|T_2| = \frac{1}{\Lambda_n}|f(X_n) - f(X_0)| \to 0,$$

as $n \to \infty$, since f is bounded and $\Lambda_n \to \infty$. Also, using Kronecker's Lemma,

$$T_1 = \frac{1}{\Lambda_n}\sum_{k=0}^{n-1}(\mathbf{E}[f(X_{k+1})|\mathcal{F}_k] - f(X_{k+1}))$$

will converge to 0 once the martingale

$$M_n^f := \sum_{k=1}^{n-1}\frac{1}{\Lambda_k}(\mathbf{E}[f(X_{k+1})|\mathcal{F}_k] - f(X_{k+1}))$$

converges a.s. Finally observing that $\mathbf{E}(f(X_{k+1})|\mathcal{F}_k)$ minimizes the L^2 distance from $f(X_{k+1})$ among $\mathcal{F}_k$ measurable square integrable random variables,

$$\begin{aligned}\mathbf{E}\langle M^f\rangle_\infty &= \sum_{k\geq 1}(\frac{1}{\Lambda_k})^2\mathbf{E}\left(f(X_{k+1})-\mathbf{E}(f(X_{k+1})|\mathcal{F}_k)\right)^2 \\ &\leq \sum_{k\geq 1}(\frac{1}{\Lambda_k})^2\mathbf{E}\left(f(X_{k+1})-f(X_k)\right)^2 \\ &\leq ||Df||_\infty \sum_{k\geq 1}(\frac{1}{\Lambda_k})^2\mathbf{E}\left(X_{k+1}-X_k\right)^2 \\ &\leq \kappa_1 \sum_{k\geq 1}\frac{\lambda_{k+1}}{\Lambda_k^2} \\ &< \infty\end{aligned}$$

for some constant κ_1, where the last inequality follows from the observation that for a positive sequence λ_k, $\sum_{k\geq 1}\lambda_{k+1}/\Lambda_k^2 < \infty$. The lemma follows. $\square$

Next we consider the limit of the first term on the right side of (6.2.16). We can regard μ_n^i to be a finite measure on the one point compactificaion of $\mathbb{R}^m$, denoted as $\bar{\mathbb{R}}^m$. In order to show that $\{\mu_n^i\}$ is a.s. a precompact sequence in $\mathcal{M}_F(\bar{\mathbb{R}}^m)$, it suffices to show that $\mu_n^i(\mathbb{R}^m)$ is an a.s. bounded sequence of $\mathbb{R}_+$ valued random variables. This is shown in the following lemma.

Lemma 6.2.8. *For $i = 1, ..., N$,*

$$\sup_n \mu_n^i(\mathbb{R}^m) < \infty, \quad a.s.$$

Proof. Let $g \in C_b^2(\mathbb{R}^m)$ be as in Lemma 6.1.1. Then for fixed $(u, v, r) \in \mathbb{R}^m \times G \times (0, \infty)$ and with notation as introduced above Lemma 6.2.6,

$$\begin{aligned}g(\mathbf{x}(1)) = g(v) &+ \int_0^1 [\nabla g(\mathbf{x}(s))\cdot(b(v)r+\sigma(v)\sqrt{r}u)]ds \\ &+ \sum_{i=1}^N \int_0^1 d_i\cdot\nabla g(\mathbf{x}(s))\alpha_i(s)d|\mathbf{y}|(s)\end{aligned} \tag{6.2.17}$$

Since $\alpha_i(s)$ is nonzero only when $\mathbf{x}(s) \in F_i$, and $\langle \nabla g(x), d_i \rangle \geq 1$, for all $x \in F_i$, $i \in \{1, ..., N\}$, we have

$$\begin{aligned}
\sum_{i=1}^{N} \mathbf{L}^{\mathbf{i}}(v,u,r) &= \sum_{i=1}^{N} \int_0^1 \alpha_i(s) d|\mathbf{y}|(s) \\
&\leq \sum_{i=1}^{N} \int_0^1 d_i \cdot \nabla g(\mathbf{x}(s)) \alpha_i(s) d|\mathbf{y}|(s) \\
&\leq |g(\mathbf{x}(1)) - g(v)| + ||\nabla g||_\infty |b(v)r + \sigma(v)\sqrt{r}u| \\
&\leq ||\nabla g||_\infty |\mathbf{x}(1) - v| + ||\nabla g||_\infty |b(v)r + \sigma(v)\sqrt{r}u| \\
&\leq ||\nabla g||_\infty (K+1)|b(v)r + \sigma(v)\sqrt{r}u|,
\end{aligned} \tag{6.2.18}$$

where the second inequality uses (6.2.17), and the last inequality uses the Lipschitz property of the Skorokhod map.

Let $\kappa_1 = ||\nabla g||_\infty (K+1)a_1$, then from (6.2.18) we have for $i \in \{1, ..., N\}$,

$$L_k^i \leq \kappa_1 \left(\sqrt{\lambda_{k+1}} |U_{k+1}| + \lambda_{k+1} \right). \tag{6.2.19}$$

Also note that, with a_1 as in Condition 6.1.3,

$$\sup_{t \in [0,1]} |x^k(t) - X_k| \leq K|b(X_k)\lambda_{k+1} + \sigma(X_k)\sqrt{\lambda_{k+1}} U_{k+1}| \leq K a_1 \sqrt{\lambda_{k+1}} |U_{k+1}| + K a_1 \lambda_{k+1}, \tag{6.2.20}$$

and for $t \in [0,1]$,

$$|\Pi_k^t - X_k| \leq t|x^k(1) - v| + (1-t)|z^k(1) - v| \leq (K+1)a_1\lambda_{k+1} + (K+1)a_1\sqrt{\lambda_{k+1}}|U_{k+1}|, \tag{6.2.21}$$

where $x^k(t) = \mathbf{x}(U_{k+1}, X_k, \lambda_{k+1}|t)$, $z^k(t) = \mathbf{z}(U_{k+1}, X_k, \lambda_{k+1}|t)$. Combining (6.2.19)-(6.2.21) we have that

$$\mathbf{E}_{X_k}(|\Pi_k^t - x^k(s_k^i)| L_k^i) \leq (2K+1)a_1\kappa_1 m \lambda_{k+1} + \varphi(\lambda_{k+1})\lambda_{k+1}, \tag{6.2.22}$$

where $\varphi : (0, \infty) \to (0, \infty)$ is a bounded function satisfying $\varphi(\alpha) \to 0$ as $\alpha \to 0$.

Next note that L_k^i is not equal to 0 only if there exists $s \in [0,1]$ such that $\alpha_i^k(s) > 0$, i.e., $x^k(s) \in F_i$, where $\alpha_i^k(t) \equiv \alpha_i(U_{k+1}, X_k, \lambda_{k+1}|t)$. And in that case (namely when $x^k(s) \in F_i$ for some $s \in [0,1]$),

$$D_i g(\Pi_k^t) \geq D_i g(x^k(s)) - ||D^2 g||_\infty |\Pi_k^t - x^k(s)| \geq 1 - ||D^2 g||_\infty |\Pi_k^t - x^k(s)|. \quad (6.2.23)$$

Let $A_k^i = \{\omega : \text{there exists } s \in [0,1] \text{ such that } \alpha_i^k(s) > 0\}$ and

$$s_k^i(\omega) = \begin{cases} \inf\{s \in [0,1] : \alpha_i^k(s) > 0\} & \text{if } \omega \in A_k^i, \\ 1 & \text{if } \omega \notin A_k^i. \end{cases}$$

Then, from (6.2.22) and (6.2.23)

$$\begin{aligned} \mathbf{E}_{X_k}[D_i g(\Pi_k^t) L_k^i 1_{A_k^i}] &\geq \mathbf{E}_{X_k}[L_k^i 1_{A_k^i}] - ||D^2 g||_\infty \mathbf{E}_{X_k}[|\Pi_k^t - x^k(s_k^i)| L_k^i 1_{A_k^i}] \\ &\geq \mathbf{E}_{X_k}[L_k^i] - ||D^2 g||_\infty ((2K+1) a_1 \kappa_1 m \lambda_{k+1} + \varphi(\lambda_{k+1}) \lambda_{k+1}) \end{aligned}$$

Thus we have

$$\begin{aligned} \langle D_i g, m_k^i \rangle &= \int_0^1 \mathbf{E}_{X_k}[D_i g(\Pi_k^t) L_k^i] dt \\ &= \int_0^1 \mathbf{E}_{X_k}[D_i g(\Pi_k^t) L_k^i 1_{A_k^i}] dt \\ &\geq \langle 1, m_k^i \rangle - ||D^2 g||_\infty \left((2K+1) a_1 m \kappa_1 \lambda_{k+1} + \varphi(\lambda_{k+1}) \lambda_{k+1}\right). \end{aligned}$$

Rearranging the terms, we have

$$\langle 1, m_k^i \rangle \leq \langle D_i g, m_k^i \rangle + ||D^2 g||_\infty \left((2K+1) a_1 m \kappa_1 \lambda_{k+1} + \varphi(\lambda_{k+1}) \lambda_{k+1}\right).$$

Summing over k from 0 to $n-1$ and i from 1 to N, we obtain

$$\sum_{i=1}^N \langle 1, \mu_n^i \rangle \leq \sum_{i=1}^N \langle D_i g, \mu_n^i \rangle + N ||D^2 g||_\infty \left((2K+1) a_1 \kappa_1 m + |\varphi|_\infty\right). \quad (6.2.25)$$

Using Lemma 6.2.6

$$\sum_{i=1}^N \mu_n^i(D_i g) = \frac{1}{\Lambda_n} \sum_{k=0}^{n-1} \mathbf{E}[g(X_{k+1}) - g(X_k) | \mathcal{F}_k] - \nu_n(\mathcal{A} g) - \xi_n^g.$$

Since $g \in C_b^2(\mathbb{R}^m)$, the second term on the right side is bounded. Also from Lemma 6.2.7, the first term converges to 0 as $n \to \infty$. Finally from Lemma 6.2.6, the third term is bounded, a.s.

From this it follows that

$$\sup_n \sum_{i=1}^N \mu_n^i(D_i g) < \infty \quad \text{a.s.}$$

Result follows on using this observation in (6.2.25). $\square$

The following lemma will be used to show that for a.e. ω, any limit point of $\mu_n^i(\omega)$ is supported on F_i, $i = 1, ..., N$.

Lemma 6.2.9. *Fix $i \in \{1, ..., N\}$. Let $\psi \in C_c^2(\mathbb{R}^m)$ be such that $\psi(x) \geq 0$ for all $x \in \mathbb{R}^m$. Suppose that there is a $\epsilon > 0$, such that $\psi(x) = 0$ if dist$(x, F_i) \leq \epsilon$. Then*

$$\int \psi(x)\mu_n^i(dx) \to 0, \text{ a.s. as } n \to \infty.$$

Proof. We have

$$\begin{aligned}
\langle \psi, \mu_n^i \rangle &= \frac{1}{\Lambda_n} \sum_{k=0}^{n-1} \langle \psi, m_k^i \rangle \\
&= \frac{1}{\Lambda_n} \sum_{k=0}^{n-1} \int_0^1 \mathbf{E}_{X_k}[\psi(\Pi_k^t) L_k^i] dt \\
&= \frac{1}{\Lambda_n} \sum_{k=0}^{n-1} \int_0^1 \mathbf{E}_{X_k} \int_0^1 \psi(\Pi_k^t) \alpha_i^k(s) d|y^k|_s dt \\
&\leq \frac{|\psi|_\infty}{\Lambda_n} \sum_{k=0}^{n-1} \mathbf{E}_{X_k} \int_0^1 \int_0^1 1_{\{|\Pi_k^t - x^k(s)| > \epsilon\}} \alpha_i^k(s) d|y^k|_s dt,
\end{aligned} \tag{6.2.26}$$

where, recall that $\alpha_i^k(s) \equiv \alpha_i(U_{k+1}, X_k, \lambda_{k+1}|s)$ and $x^k(s) = \mathbf{x}(U_{k+1}, X_k, \lambda_{k+1}|s)$, $y^k(s) = \mathbf{y}(U_{k+1}, X_k, \lambda_{k+1}|s)$. The last inequality in the above expression follows from the fact that $\alpha_i^k(s) > 0$ only when $x^k(s) \in F_i$ and if for such a s, $|\Pi_k^t - x^k(t)| \leq \epsilon$,we have by our choice of ψ that $\psi(\Pi_k^t) = 0$.

Next note that

$$\begin{aligned}
\{(t, s, \omega) :& |\Pi_k^t - x^k(s)| > \epsilon\} \\
&\subset \{(t, s, \omega) : |x^k(1) - x^k(s)| > \epsilon\} \cup \{(t, s, \omega) : |z^k(1) - x^k(s)| > \epsilon\},
\end{aligned}$$

where $z^k(t)$ is as introduced below (6.2.20), namely $z^k(t) = \mathbf{z}(U_{k+1}, X_k, \lambda_{k+1}|t)$.

Also, from the Lipschitz property of the Skorokhod map,

$$|x^k(1) - x^k(s)| \leq Ka_1\lambda_{k+1} + Ka_1\sqrt{\lambda_{k+1}}|U_{k+1}|,$$

and

$$|z^k(1)-x^k(s)| \leq |z^k(1)-X_k|+|X_k-x^k(s)| \leq (K+1)a_1\lambda_{k+1}+(K+1)a_1\sqrt{\lambda_{k+1}}|U_{k+1}|.$$

Thus

$$\{\omega : |\Pi_k^t - x^k(s)| > \epsilon \text{ for some } t, s \in [0,1]\} \subset \{\omega : |U_{k+1}(\omega)| \geq p_k\},$$

where $p_k = \frac{\epsilon/((K+1)a_1)-\lambda_{k+1}}{\sqrt{\lambda_{k+1}}}$. Using this observation in (6.2.26), we have

$$\begin{aligned}
\langle \psi, \mu_n^i \rangle &\leq \frac{|\psi|_\infty}{\Lambda_n} \sum_{k=0}^{n-1} \mathbf{E}_{X_k} \int_0^1 \int_0^1 1_{\{|U| \geq p_k\}} \alpha_i^k(s) d|y^k|_s dt \\
&\leq \frac{|\psi|_\infty}{\Lambda_n} \sum_{k=0}^{n-1} \mathbf{E}_{X_k} \left(1_{\{|U| \geq p_k\}} \int_0^1 \alpha_i^k(s) d|y^k|_s \right) \\
&\leq \frac{|\psi|_\infty}{\Lambda_n} \sum_{k=0}^{n-1} \sqrt{\left(\mathbf{E}_{X_k} \left(\int_0^1 \alpha_i^k(s) d|y^k|_s \right)^2 \right) \mathbb{P}(|U| \geq p_k)}.
\end{aligned}$$

From (6.2.18) it follows that for some $\kappa_1 \in (0, \infty)$, $\sup_k \mathbf{E}_{X_k} \left(\int_0^1 \alpha_i^k(s) d|y^k|_s \right)^2 \leq \kappa_1$. Also using Condition 6.1.6, $\mathbf{E}|U|^j < \infty$ for all $j \geq 1$. Choose k_0 large enough so that $\lambda_{k+1} \leq \frac{\epsilon}{2(K+1)a_1}$ for all $k \geq k_0$.

Fix $j > 4$, then

$$\langle \psi, \mu_n^i \rangle \leq \frac{|\psi|_\infty}{\Lambda_n} \sqrt{\kappa_1} k_0 + \frac{|\psi|_\infty}{\Lambda_n} \sum_{k=k_0}^{n-1} \sqrt{\kappa_1} (\mathbf{E}|U|^j)^{1/2} p_k^{-j/2}.$$

The result now follows on observing that for some $\kappa_2 \in (0, \infty)$, $p_k^{-j/2} \leq \kappa_2 \lambda_{k+1}^{j/4}$ for all $k \geq k_0$. □

We are now ready to complete the proof of Theorem 6.1.3.

Proof of Theorem 6.1.3. Fix $f \in C_c^2(G)$. Then such a function can be extended to a function in $C_c^2(\mathbb{R}^m)$. We denote this function once more by f. Then from Lemma 6.2.6,

$$\frac{1}{\Lambda_n}\sum_{k=0}^{n-1}\mathbf{E}[f(X_{k+1}) - f(X_k)|\mathcal{F}_k] = \sum_{i=1}^{N}\mu_n^i(D_i f) + \nu_n(\mathcal{A}f) + \xi_n^f. \tag{6.2.27}$$

From Lemmas 6.2.5, 6.2.6, 6.2.7 and 6.2.8, there exists $\Omega_0 \in \mathcal{F}$ such that $\mathbb{P}(\Omega_0) = 1$ and for every $\omega \in \Omega_0$,

- $\{\nu_n(\omega)\}_n$ is precompact in $\mathcal{P}(G)$,
- $\{\mu_n^i(\omega)\}_n$ is precompact in $\mathcal{M}_F(\bar{\mathbb{R}}^m)$, for every $i = 1, ..., N$,
- Left hand side of (6.2.27) converges to 0,
- $\xi_n^f(\omega)$ converges to 0.

Fix a $\omega \in \Omega_0$ and let $\nu_\infty(\omega)$, $\mu_\infty^i(\omega)$, $i = 1, ..., N$, be a subsequential limit of $\nu_n(\omega)$ and $\mu_n^i(\omega)$, respectively. Then from (6.2.27) and the above observations, we have (suppressing ω)

$$\nu_\infty(\mathcal{A}f) + \sum_{i=1}^{N}\mu_\infty^i(D_i f) = 0.$$

To complete the proof, in view of Theorem 6.2.1, it suffices to argue that

$$\int_{\mathbb{R}^m} 1_{F_i^c}(x)\mu_\infty^i(\omega)(dx) = 0. \tag{6.2.28}$$

By convergence of μ_n^i to μ_∞^i, we have for every ψ as in Lemma 6.2.9,

$$\int_{\mathbb{R}^m} \psi(x)\mu_\infty^i(\omega)(dx) = 0.$$

Therefore

$$\int_{\mathbb{R}^m} 1_{F_i^{\epsilon,r}}(x)\mu_\infty^i(dx) = 0 \quad \forall \epsilon, r > 0,$$

where $F_i^{\epsilon,r} = \{x \in \mathbb{R}^m | \operatorname{dist}(x, F_i) \geq \epsilon \text{ and } |x| \leq r\}$. The equality in (6.2.28) now follows on sending $\epsilon \to 0$ and $r \to \infty$. $\square$

6.2.3 Proof of Theorem 6.1.4.

Recall c from Lemma 6.2.1 and ϖ from (6.2.4). Fix $\zeta \in (0, \varpi c)$. We will prove the theorem with such a choice of ζ. Consider an f as in the statement of the theorem. Then there exists constant κ_1 such that $|f(x)| \le \kappa_1 e^{\zeta|x|}$ Without loss of generality, we assume $f \ge 0$.

From Theorem 6.1.3, for any $L > 0$, we have

$$\int (f \wedge L) d\nu_n \to \int (f \wedge L) d\nu \quad \text{a.s.}$$

In order to prove the theorem, it suffices to show that

$$\lim_{L\to\infty} \sup_n \left[\int f d\nu_n - \int (f \wedge L) d\nu_n \right] = 0$$

and

$$\lim_{L\to\infty} \left[\int f d\nu - \int (f \wedge L) d\nu \right] = 0.$$

First, consider

$$\begin{aligned} \sup_n \left[\int f d\nu_n - \int (f \wedge L) d\nu_n \right] &\le \sup_n \int 1_{f>L} f d\nu_n \\ &\le \sup_n \left(\nu_n^{1/p}(f > L) [\nu_n(f^q)]^{1/q} \right), \end{aligned}$$

where $p, q \in (1, \infty)$ are such that $p^{-1} + q^{-1} = 1$ and the last inequality follows from Hölder's inequality. Choose $q > 1$ such that $\zeta q < \varpi c$, then from Lemma 6.2.5 we have

$$\sup_n [\nu_n(f^q)]^{1/q} \le \kappa_1 \sup_n [\int e^{\zeta q|x|} \nu_n(dx)]^{1/q} \le \kappa_1 \sup_n \nu_n^{1/q}(V) < \infty, \quad \text{a.s.} \tag{6.2.29}$$

Using Markov's Inequality, we have

$$\nu_n^{1/p}(f > L) \le \frac{\nu_n^{1/p}(f)}{L^{1/p}},$$

which using (6.2.29) converges to 0 as L goes to infinity. Combining the above three displays, we have

$$\sup_n \left[\int f d\nu_n - \int (f \wedge L) d\nu_n \right] \to 0, \quad \text{a.s. as } L \to \infty. \tag{6.2.30}$$

Also, from Fatou's lemma we have, for a.e. ω,

$$\begin{aligned} \int f d\nu - \int (f \wedge L) d\nu &= \int (f - f \wedge L) d\nu \\ &\leq \liminf_n \int (f - f \wedge L) d\nu_n \\ &\leq \sup_n \int (f - f \wedge L) d\nu_n. \end{aligned}$$

Using (6.2.30) the last expression converges to 0 as $L \to \infty$. The result follows.

6.3 Proof of Theorem 6.1.5.

We begin with a few preliminary lemmas.

Lemma 6.3.1. *If $\phi \in C^2(G)$, then*

$$\Lambda_n \nu_n(\mathcal{A}\phi) = \sum_{k=1}^n \lambda_k \mathcal{A}\phi(X_{k-1}) = Z_n^{(0)} - (N_n + \sum_{i=1}^4 Z_n^{(i)} + \sum_{i=1}^4 Y_n^{(i)})$$

with

$$\begin{aligned} Z_n^{(0)} &= \phi(X_n) - \phi(X_0), \\ N_n &= \sum_{k=1}^n \sqrt{\lambda_k} \langle \nabla\phi(X_{k-1}), \sigma(X_{k-1})U_k \rangle, \\ Z_n^{(1)} &= \frac{1}{2} \sum_{k=1}^n \lambda_k^2 b(X_{k-1})^T D^2\phi(X_{k-1}) b(X_{k-1}), \\ Z_n^{(2)} &= \sum_{k=1}^n \lambda_k^{3/2} b(X_{k-1})^T D^2\phi(X_{k-1}) \sigma(X_{k-1}) U_k, \\ Z_n^{(3)} &= \frac{1}{2} \sum_{k=1}^n \lambda_k [(\sigma(X_{k-1})U_k)^T D^2\phi(X_{k-1})(\sigma(X_{k-1})U_k) \\ &\quad -\mathbf{E}((\sigma(X_{k-1})U_k)^T D^2\phi(X_{k-1})(\sigma(X_{k-1})U_k)|\mathcal{F}_{k-1})], \\ Z_n^{(4)} &= \sum_{k=1}^n R_2(X_{k-1}, X_k), \end{aligned}$$

and

$$\begin{aligned}
Y_n^{(1)} &= \sum_{k=1}^{n} \langle \nabla\phi(X_{k-1}), y_{k-1}\rangle, \\
Y_n^{(2)} &= \frac{1}{2}\sum_{k=1}^{n} y_{k-1}^T D^2\phi(X_{k-1}) y_{k-1}, \\
Y_n^{(3)} &= \sum_{k=1}^{n} \lambda_k b(X_{k-1})^T D^2\phi(X_{k-1}) y_{k-1}, \\
Y_n^{(4)} &= \sum_{k=1}^{n} \sqrt{\lambda_k} y_{k-1}^T D^2\phi(X_{k-1})\sigma(X_{k-1})U_k,
\end{aligned}$$

where $R_2(x,y) = \phi(y) - \phi(x) - \langle\nabla\phi(x), y-x\rangle - \frac{1}{2}(y-x)^T D^2\phi(x)(y-x)$, *and* $y_k = \mathbf{y}(U_{k+1}, X_k, \lambda_{k+1}|1)$.

Proof. Denote $\delta\phi(X_k) = \phi(X_k) - \phi(X_{k-1})$ and $\delta X_k = X_k - X_{k-1}$. We deduce from (6.1.5) that

$$\begin{aligned}
\delta\phi(X_k) =& \langle\nabla\phi(X_{k-1}), \delta X_k\rangle + \frac{1}{2}\delta X_k^T D^2\phi(X_{k-1})\delta X_k + R_2(X_{k-1}, X_k) \\
=& \langle\nabla\phi(X_{k-1}), y_{k-1}\rangle + \lambda_k \mathcal{A}\phi(X_{k-1}) + \sqrt{\lambda_k}\langle\nabla\phi(X_{k-1}), \sigma(X_{k-1})U_k\rangle \\
&+ \frac{1}{2} y_{k-1}^T D^2\phi(X_{k-1}) y_{k-1} + \frac{1}{2}\lambda_k^2 b(X_{k-1})^T D^2\phi(X_{k-1}) b(X_{k-1}) \\
&+ \frac{1}{2}\lambda_k \big[(\sigma(X_{k-1})U_k)^T D^2\phi(X_{k-1})(\sigma(X_{k-1})U_k) \\
&\quad - \mathbf{E}((\sigma(X_{k-1})U_k)^T D^2\phi(X_{k-1})(\sigma(X_{k-1})U_k)|\mathcal{F}_{k-1})\big] \\
&+ \lambda_k b(X_{k-1})^T D^2\phi(X_{k-1}) y_{k-1} + \lambda_k^{3/2} b(X_{k-1})^T D^2\phi(X_{k-1})\sigma(X_{k-1})U_k \\
&+ \sqrt{\lambda_k} y_{k-1}^T D^2\phi(X_{k-1})\sigma(X_{k-1})U_k + R_2(X_{k-1}, X_k).
\end{aligned}$$

The lemma follows by summing the above equality over $k = 1, ..., n$ and rearranging the terms. □

Lemma 6.3.2. *Let* $W : G \to \mathbb{R}$ *be a continuous function such that* $\sup_{n\in\mathbb{N}} \nu_n(W) < \infty$, *a.s. Let* $\phi \in C^1(G)$, *be such that* $\lim_{|x|\to\infty} |\nabla\phi(x)|^2/W(x) = 0$. *Then*

$$\frac{1}{\sqrt{\Lambda_n}}\sum_{k=1}^{n}\sqrt{\lambda_k}\langle\nabla\phi(X_{k-1}), \sigma(X_{k-1})U_k\rangle \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_G |\sigma^T\nabla\phi|^2 d\nu\right).$$

Proof. This lemma follows from Theorem 6.1.3 using the martingale central limit theorem, along the lines of Proposition 2 of [57]. Details are left to the reader. $\square$

Lemma 6.3.3. *Under the assumptions of Theorem 6.1.5(b), we have,*

$$\frac{Z_n^{(4)}}{\Lambda_n^{(3/2)}} \xrightarrow{\mathbb{P}} \frac{1}{6}\int_G \int_{\mathbb{R}^m} D^3\phi(x)(\sigma(x)u)^{\otimes 3}\mu(du)\nu(dx),$$

as $n \to \infty$.

Proof. The proof is similar to that of Lemma 10 of [57] except for the treatment of reflection terms. Using the notation above Theorem 6.1.5 and in Lemma 6.3.1, we have

$$R_2(x,y) = \frac{1}{6}D^3\phi(x)(y-x)^{\otimes 3} + R_4(x,y), \tag{6.3.1}$$

with

$$|R_4(x,y)| \le \frac{L}{6}|y-x|^4,$$

where L is the Lipschitz constant for $D^3\phi$. Hence

$$R_2(X_{k-1},X_k) = \frac{1}{6}D^3\phi(X_{k-1})(\delta X_k)^{\otimes 3} + r_k, \tag{6.3.2}$$

with

$$|r_k| \le \frac{L}{6}|\delta X_k|^4 \le \kappa_1 \lambda_k^2(1+|U_k|^4), \quad k \in \mathbb{N},$$

for some $\kappa_1 \in (0,\infty)$. Since $\mathbf{E}|U_k|^4 := \mu_4 < \infty$ from Condition 6.1.6, we have

$$\mathbf{E}\sum_{k=1}^n |r_k| \le \kappa_1(1+\mu_4)\sum_{k=1}^n \lambda_k^2.$$

From the assumption $\lim_{n\to\infty}(1/\sqrt{\Lambda_n})\sum_{k=1}^n \lambda_k^{3/2} = \tilde{\lambda} \in (0,+\infty]$, we deduce that $\lim_{n\to\infty}\sum_{k=1}^n \lambda_k^{3/2} = +\infty$ and

$$\lim_{n\to\infty}\sum_{k=1}^n \lambda_k^2/\Lambda_n^{(3/2)} = 0. \tag{6.3.3}$$

Therefore,

$$\frac{1}{\Lambda_n^{(3/2)}}\sum_{k=1}^n r_k \xrightarrow{L^1} 0. \tag{6.3.4}$$

Now consider the first term on the right side of (6.3.2).

$$\begin{aligned}
&D^3\phi(X_{k-1})(\delta X_k)^{\otimes 3}\\
&\qquad = D^3\phi(X_{k-1})(\lambda_k b(X_{k-1}) + \sqrt{\lambda_k}\sigma(X_{k-1})U_k + y_{k-1})^{\otimes 3}\\
&\qquad = \lambda_k^{3/2} D^3\phi(X_{k-1})(\sqrt{\lambda_k} b(X_{k-1}) + \sigma(X_{k-1})U_k)^{\otimes 3} + f_k^{(1)}(X_{k-1}, U_k)\\
&\qquad = \lambda_k^{3/2} D^3\phi(X_{k-1})(\sigma(X_{k-1})U_k)^{\otimes 3} + f_k^{(2)}(X_{k-1}, U_k) + f_k^{(1)}(X_{k-1}, U_k),
\end{aligned}$$

where $f_k^{(1)}(X_{k-1}, U_k)$ and $f_k^{(2)}(X_{k-1}, U_k)$ are defined through the second and third equalities, respectively.

Next observe that

- From the assumptions, we have

$$|\lambda_k b(X_{k-1}) + \sqrt{\lambda_k}\sigma(X_{k-1})U_k| \le a_1\sqrt{\lambda_k}(|U_k| + \sqrt{\lambda_0}).$$

- From (6.2.14) and (6.2.19), we have $y_{k-1} = \sum_{i=1}^N d_i L_{k-1}^i$ and for some $\kappa_2 \in (0, \infty)$,

$$L_{k-1}^i \le \kappa_2\sqrt{\lambda_k}(|U_k| + 1), \text{ for all } k \in \mathbb{N}$$

- The term L_{k-1}^i is non zero only if there exists $s \in [0, 1]$ such that $x_{k-1}(s) \in F_i$, where $x_{k-1}(s) = \mathbf{x}(U_k, X_{k-1}, \lambda_k|s)$. And in that case, we have from (6.2.20), the Lipschitz property of $D^3\phi$ and (6.1.8) that, for some $\kappa_3 \in (0, \infty)$,

$$|D^3_{\cdot jk}\phi(X_{k-1}) \cdot d_i| \le \kappa_3\sqrt{\lambda_k}(|U_k| + 1), \quad \forall j, k.$$

Combining these estimates, we see that $\mathbf{E}\sum_{k=1}^n |f_k^{(1)}(X_{k-1}, U_k)| \le \kappa_4 \sum_{k=1}^n \lambda_k^2$. Using (6.3.3) we now have

$$\frac{1}{\Lambda_n^{(3/2)}} \sum_{k=1}^n f_k^{(1)}(X_{k-1}, U_k) \xrightarrow{L^1} 0. \tag{6.3.5}$$

For the term $f_k^{(2)}(X_{k-1},U_k)$, using the boundedness of $D^3\phi$, b, and σ, it can be easily checked that $\mathbf{E}|f_k^{(2)}(X_{k-1},U_k)| \le \kappa_5\lambda_k^2$. Thus

$$\mathbf{E}\sum_{k=1}^{n}|f_b(X_{k-1},U_k)| \le \kappa_5\sum_{k=1}^{n}\lambda_k^2,$$

and so using (6.3.3) once again we have

$$\frac{1}{\Lambda_n^{(3/2)}}\sum_{k=1}^{n} f_b(X_{k-1},U_k) \xrightarrow{\mathbb{P}} 0. \tag{6.3.6}$$

Let $\Theta(X_{k-1},U_k) = D^3\phi(X_{k-1})(\sigma(X_{k-1})U_k)^{\otimes 3}$. Since $\sup_k \mathbf{E}|\Theta(X_{k-1},U_k)|^2 < \infty$ and $\lim_{n\to\infty}\Lambda_n^{(3)}/(\Lambda_n^{(3/2)})^2 = 0$, we have

$$\frac{1}{\Lambda_n^{(3/2)}}\sum_{k=1}^{n}\lambda_k^{3/2}[\Theta(X_{k-1},U_k) - \mathbf{E}(\Theta(X_{k-1},U_k)|\mathcal{F}_{k-1})] \xrightarrow{L^2} 0. \tag{6.3.7}$$

Observe that $\mathbf{E}(\Theta(X_{k-1},U_k)|\mathcal{F}_{k-1}) = J(X_{k-1})$, where J is given by

$$J(x) := \int_{\mathbb{R}^m} D^3\phi(x)(\sigma(x)u)^{\otimes 3}\mu(du).$$

Since $\Lambda_n^{(3/2)} \to \infty$ as $n\to\infty$, we can apply Theorem 6.1.3 to the measure $\tilde{\nu}_n = \frac{1}{\Lambda_n^{(3/2)}}\sum_{k=1}^{n}\lambda_k^{3/2}\delta_{X_{k-1}}$. Since J is continuous and bounded, we have $\lim_{n\to\infty}\tilde{\nu}_n(J) = \int J d\nu$ a.s., and the lemma follows on combining this fact with (6.3.1)-(6.3.7). $\square$

We are now ready to prove Theorem 6.1.5.

Proof of Theorem 6.1.5. The proof is similar as the proof of Theorem 9 of [57], once again the main difference is in the treatment of reflection terms. Using the notation of Lemma 6.3.1, we first observe that, for any sequence of positive numbers $\{a_n\}_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} a_n = \infty$, we have $Z_n^{(0)}/a_n \to 0$ in probability. This is because, from Lemma 6.2.4, the sequence $\{X_n\}_{n\in\mathbb{N}}$ is tight, and consequently so is $\{\phi(X_n)\}_{n\in\mathbb{N}}$ as well.

We also derive from the definitions of $Z_n^{(1)}$, $Z_n^{(2)}$ and $Z_n^{(3)}$ the inequalities

$$\mathbf{E}|Z_n^{(1)}| \le \kappa_1 \sum_{k=1}^{n} \lambda_k^2 ||D^2\phi||_\infty, \tag{6.3.8}$$

and

$$\mathbf{E}|Z_n^{(i)}|^2 \le \kappa_1 \sum_{k=1}^{n} \lambda_k^2 ||D^2\phi||_\infty^2, \quad i = 2, 3, \tag{6.3.9}$$

for some $\kappa_1 \in (0, \infty)$, for all $n \ge 1$.

(a) Assume that $\lim_{n\to\infty}(1/\sqrt{\Lambda_n})\Lambda_n^{(3/2)} = 0$. We have $\lim_{n\to\infty} \sum_{k=1}^n \lambda_k^2/\sqrt{\Lambda_n} = 0$, and it follows from (6.3.8) that $Z_n^{(1)}/\sqrt{\Lambda_n} \xrightarrow{L^1} 0$. We also deduce from (6.3.9), that $Z_n^{(j)}/\sqrt{\Lambda_n} \xrightarrow{L^2} 0$, for $j = 2, 3$. Consider now $Z_n^{(4)}$. Denoting the Lipschitz norm of $D^2\phi$ by L, we have

$$|R_2(X_{k-1}, X_k)| \le \frac{L}{2}|\Delta X_k|^3 \le \frac{L}{2} a_1^3 K^3 (\lambda_k + \sqrt{\lambda_k}|U_k|)^3,$$

where the second inequality follows from the Lipschitz property of the Skorokhod map (Condition 6.1.1). Thus, there exists $\kappa_2 \in (0, \infty)$ such that, for all $n \ge 1$,

$$\mathbf{E}|Z_n^{(4)}| \le \kappa_2 \sum_{k=1}^{n} \lambda_k^{3/2}, \tag{6.3.10}$$

and therefore $Z_n^{(4)}/\sqrt{\Lambda_n} \xrightarrow{L^1} 0$.

We now, consider $Y_n^{(j)}$, for $j = 1, 2, 3, 4$.

$$Y_n^{(1)} = \sum_{k=1}^{n} \langle \nabla\phi(X_{k-1}), y_{k-1} \rangle = \sum_{k=1}^{n} D_i\phi(X_{k-1}) L_{k-1}^i.$$

From (6.2.19), we have $|L_{k-1}^i| \le \kappa_3 \sqrt{\lambda_k}(|U_k| + 1)$. Also, for any fixed i, L_{k-1}^i is not equal to 0 only if there exists $x \in F_i$, such that $||X_{k-1} - x|| \le a_1 K \lambda_k + a_1 K \sqrt{\lambda_k}|U_k|$; and in that case, using Taylor's theorem and the Lipschitz property of $D^2\phi$, there exists $\kappa_4 \in (0, \infty)$, such that,

$$|D_i\phi(X_{k-1}) - D_i\phi(x) - (X_{k-1} - x)^T D^2\phi(x) d_i| \le \kappa_4 ||X_{k-1} - x||^2.$$

Combining this with (6.1.7), we have

$$|D_i\phi(X_{k-1})| \leq \kappa_4 ||X_{k-1} - x||^2.$$

Thus we have

$$\mathbf{E}|Y_n^{(1)}| \leq \kappa_5 \sum_{k=1}^{n} \lambda_k^{3/2}, \tag{6.3.11}$$

for some constant κ_5. Using similar arguments as above, we obtain:

$$\mathbf{E}|Y_n^{(j)}| \leq \kappa_5 \sum_{k=1}^{n} \lambda_k^{3/2}, \quad j = 2, 3, 4. \tag{6.3.12}$$

Thus we have that $Y_n^{(j)}/\sqrt{\Lambda_n} \xrightarrow{L^1} 0$, for $j = 1, 2, 3, 4$.

From Lemma 6.2.5, and recalling the definition of V (see (6.2.1)) we have that, for every $\zeta \in (0, c\varpi)$,

$$\sup_{n\in\mathbb{N}} \int_G e^{\zeta|x|}\nu_n(dx) < \infty, \quad \text{a.s.}$$

For such a ζ, under the assumption that $\lim_{|x|\to\infty} e^{-\zeta|x|}|\nabla\phi(x)|^2 = 0$, applying Lemma 6.3.2, we now have

$$\frac{N_n}{\sqrt{\Lambda_n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_G |\sigma^T\nabla\phi|^2 d\nu\right).$$

This completes the proof of part (a).

(b) Assume now that $\lim_{n\to\infty}(1/\sqrt{\Lambda_n})\Lambda_n^{(3/2)} = \tilde{\lambda} \in (0, +\infty]$. We then have that

$$\lim_{n\to\infty} \Lambda_n^{(3/2)} = +\infty \text{ and } \lim_{n\to\infty} \sum_{k=1}^{n} \lambda_k^2/\Lambda_n^{(3/2)} = 0.$$

As before, $Z_n^{(0)}/\Lambda_n^{(3/2)} \xrightarrow{\mathbb{P}} 0$. It follows from (6.3.8) that $Z_n^{(1)}/\Lambda_n^{(3/2)} \xrightarrow{L^1} 0$, and from (6.3.9) that $Z_n^{(j)}/\Lambda_n^{(3/2)} \xrightarrow{L^2} 0$, for $j = 2, 3$.

Under the assumptions of part (b) (i.e. that $D^3\phi$ is bounded, Lipschitz and (6.1.8) holds), we have, using similar arguments as in part (a), for some $\kappa_6 \in (0, \infty)$,

$$\mathbf{E}|Y_n^{(j)}| \leq \kappa_6 \sum_{k=1}^{n} \lambda_k^2, \quad j = 1, ..., 4;\ n \geq 1. \tag{6.3.13}$$

Thus we have that $Y_n^{(j)}/\Lambda_n^{(3/2)} \xrightarrow{L^1} 0$, for $j = 1, 2, 3, 4$.

Applying Lemma 6.3.2 once again, we have, for ϕ satisfying $\lim_{|x|\to\infty} e^{-\zeta|x|}|\nabla\phi(x)|^2 = 0$,

$$\frac{N_n}{\sqrt{\Lambda_n}} \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \int_G |\sigma^T \nabla\phi|^2 d\nu\right). \tag{6.3.14}$$

Also from Lemma 6.3.3

$$\frac{Z_n^{(4)}}{\Lambda_n^{(3/2)}} \xrightarrow{\mathbb{P}} \frac{1}{6}\int_G \int_{\mathbb{R}^m} D^3\phi(x)(\sigma(x)u)^{\otimes 3}\mu(du)\nu(dx) = -\tilde{m}. \tag{6.3.15}$$

Now, if $\tilde{\lambda} < +\infty$, we have from the above observations that $Z_n^{(j)}/\sqrt{\Lambda_n} \xrightarrow{\mathbb{P}} 0$, for $j = 0, 1, 2, 3$, $Y_n^{(j)}/\sqrt{\Lambda_n} \xrightarrow{\mathbb{P}} 0$, for $j = 1, 2, 3, 4$ and

$$\frac{Z_n^{(4)}}{\sqrt{\Lambda_n}} \xrightarrow{\mathbb{P}} -\tilde{\lambda}\tilde{m}. \tag{6.3.16}$$

The statement in (6.1.9) now follows on combining this with (6.3.14).

Finally, if $\tilde{\lambda} = +\infty$, we have $Z_n^{(j)}/\Lambda_n^{(3/2)} \xrightarrow{\mathbb{P}} 0$, for $j = 0, 1, 2, 3$, $Y_n^{(j)}/\Lambda_n^{(3/2)} \xrightarrow{\mathbb{P}} 0$, for $j = 1, 2, 3, 4$ and $N_n/\Lambda_n^{(3/2)} \xrightarrow{\mathbb{P}} 0$, and (6.1.10) follows from (6.3.15). This completes the proof of Theorem 6.1.5. □

6.4 Numerical Results.

6.4.1 Evaluation of the Euler Time Step.

A key step in simulating the sequence $\{X_k\}$ in (6.1.5) is the evaluation of $\mathcal{S}(X_k, Y_{k+1} - X_k)$, where $\mathcal{S} : G \times \mathbb{R}^m \to G$ is the time-1 Skorokhod map defined in (6.1.4). In this section we describe a procedure for computing $\mathcal{S}(x, v)$ that uses well known relationships between Skorokhod problems and linear complementary problems (LCP). We restrict ourselves to a setting where $N = m$ and $G = \mathbb{R}_+^m$. We begin by recalling the basic formulation of the LCP (see [24]). For $j \in \mathbb{N}$, a $j \times j$ matrix R and a j-dimensional vector θ, the LCP for (R, θ) is to find vectors $u, v \in \mathbb{R}^j$ such that

$$\begin{cases} u \geq 0, v \geq 0; \\ v = \theta + Ru; \\ u \cdot v = 0. \end{cases}$$

It is well known (see [35] and [13]) that with $R = [d_1, ..., d_m]$, under Condition 6.1.1, for every $\theta \in \mathbb{R}^m$, the LCP for (R, θ) admits a unique solution $(u, v) \equiv (\mathcal{L}^1_m(R, \theta), \mathcal{L}^2_m(R, \theta))$, and furthermore $\mathcal{L}^2_m(R, \theta) = \mathcal{S}(0, \theta)$. Thus the evaluation of $\mathcal{S}(0, \theta)$ reduces to solving the above LCP for which numerous algorithms are available. In the examples considered in the current work we used a quadratic programming algorithm. Evaluation of $\mathcal{S}(x, \theta)$ for $x \neq 0$ can be carried out using a localization procedure as follows.

Fix $x \in G$ and let $J = \text{In}(x) = \{j \in \{1, ..., m\} | \langle x, e_j \rangle = 0\}$. Let $P_J = \{z \in \mathbb{R}^m | \langle z, e_j \rangle = 0, \ \forall j \in J^c\}$. Let $\pi_J : \mathbb{R}^m \to P_J$ be the orthogonal projection:

$$\pi_J(z) = z - \sum_{j \in J^c} \langle z, e_j \rangle e_j.$$

Suppose that $|J| = p$ and $J = \{i_1, ..., i_p\}$. Define a $p \times p$ matrix R_J be the relation $R_J(k, l) = (\pi_J d_{i_l})_{i_k}$, for $k, l = 1, ...p$. Let $u_J, v_J \in \mathbb{R}^p$ be the solution of LCP for $(R_J, \pi_J \theta)$, i.e., $(u_J, v_J) = (\mathcal{L}^1_p(R_J, \pi_J \theta), \mathcal{L}^2_p(R_J, \pi_J \theta))$. Once again unique solvability of LCP for $(R_J, \pi_J \theta)$ is assured from Condition 6.1.1. Denote $u_J = (\eta_1, ..., \eta_p)$ and define $x_1(t) = x + \theta t + t \sum_{j=1}^p \eta_j d_{i_j}$. Let

$$\tau_1 = \inf\{t \geq 0 | \text{In}(x_1(t)) \neq \text{In}(x)\}.$$

We define $\tau_1 = \infty$ if the above set is empty. Then $\Gamma(x + \theta i)(t) = x_1(t)$ for all $t < \tau_1$. If $\tau_1 < \infty$ set the initial point to be $x_1 = x_1(\tau_1)$ and define the trajectory $\{x_2(t)\}_{t \geq 0}$ in a similar way as $\{x_1(t)\}$ by replacing x with x_1. Set $\tau_2 = \inf\{t \geq 0 | \text{In}(x_2(t)) \neq \text{In}(x_1)\}$.

Then

$$\Gamma(x+\theta i)(\tau_1+t) = \Gamma(x_1+\theta i)(t) = x_2(t), \text{ for all } t < \tau_2.$$

Define now recursively trajectory $\{x_j(t)\}$ with time points τ_j and end points $x_j(\tau_j)$, $j = 3, 4, ...$ Let j_0 be such that $\sum_{i=1}^{j_0} \tau_i < 1 \leq \sum_{i=1}^{j_0+1} \tau_i$. Then

$$\mathcal{S}(x,\theta) = \Gamma(x+\theta i)(1) = \Gamma(x_{j_0}+\theta i)(1-\sum_{i=1}^{j_0} \tau_i).$$

Thus the evaluation of $\mathcal{S}(x,\theta)$ can be carried out by recursively solving a sequence of LCP problems.

One difficulty in implementing the above scheme is the possibility that $\sum_{i=1}^{\infty} \tau_i \leq 1$. However using regularity property of the Skorokhod map, we see that this occurs only when $\mathcal{S}(x,\theta)$ is zero. Thus in the practical implementation of the algorithm we fix a finite threshold L and carry out the above recursive procedure at most L times and set $\mathcal{S}(x,\theta) = 0$ if $\sum_{i=1}^{L} \tau_i < 1$.

6.4.2 Results.

A 3-d Example with Product Form Stationary Distribution.

First, we consider a 3-d example, where the reflection matrix is of the form: $R = I + Q$, where I is the identity matrix, and Q is given as the following:

$$Q = \begin{bmatrix} 0 & 0.1 & -0.2 \\ -0.1 & 0 & 0 \\ 0.2 & 0 & 0 \end{bmatrix}.$$

The spectral radius of Q is less than 1, so the conditions in this paper holds. Take the drift function $b(x) = [-1/2, -1/2, -1/2]^T$ and $\sigma(x) = I$, for all $x \in \mathbb{R}_+^3$. According to [47], the stationary distribution in this case is of product form: $\exp(1.1667) \otimes$

exp(1.0938) $\otimes$ exp(0.8537), where exp(μ) is the exponential distribution with mean μ. We set our initial point to $X_0 = [1,1,1]^T$, and simulate $\{X_k\}_{k=1}^n$ defined by equation (6.1.5), taking $U_k \sim \mathcal{N}(0, I)$, $\lambda_k = 1/\sqrt{k}$ and $n = 10^7$. Figure 6.1 shows the comparison between the exact distribution with the first-coordinate marginal of the measure ν_n using the last 10^6 simulated trajectory points. Note that the actual time span for 10^7 steps with $\alpha = 0.5$ is about 6323, and the computation time is $2383.63s$.[1]

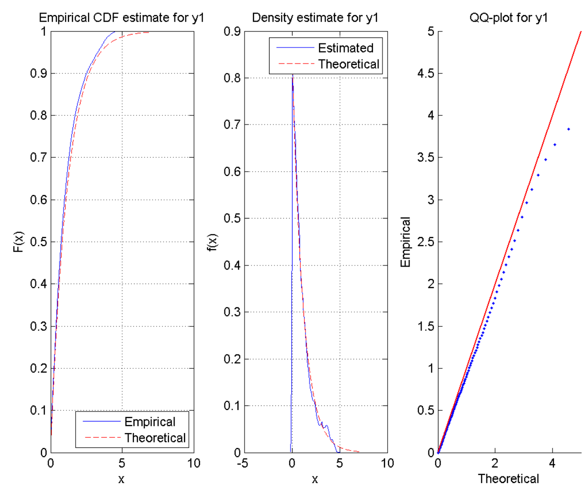

Figure 6.1: Comparison between the sample distribution with theoretical distribution. The left one is the comparison between the empirical cumulative distribution function (cdf) and the theoretical cdf. The middle one is the the comparison between the estimated density function and the theoretical density function. And the right one is the qq-plot of the empirical quantiles versus the theoretical quantiles.

In this example, we summarize the value of $\Lambda_n^r(\int x_1 \nu_n(dx) - \text{mean})$ with different rate r evaluated for different n in Table 6.1. In Theorem 6.1.5, we prove a convergence rate result for certain class of functions. Although the function x_1 is not in this class, we could still get the convergence. And Table 6.1 will give some idea about the convergence rate.

[1]The computing time in this paper is measured in Linux system with 2.93GHz Intel processor with algorithm implemented in Matlab.

Table 6.1: Value of $\Lambda_n^r(\int x_1\nu_n(dx) - \text{mean})$ with different rate r evaluated for different n.

$r\backslash n$	10	10^2	10^3	10^4	10^5	10^6	10^7
0.0000	0.2725	-0.4576	-0.2598	-0.0418	-0.0799	-0.0611	3.4349e-5
0.1000	0.3203	-0.6129	-0.3924	-0.0710	-0.1523	-0.1306	8.2415e-5
0.2000	0.3763	-0.8209	-0.5928	-0.1205	-0.2903	-0.2793	1.9774e-4
0.3000	0.4423	-1.0996	-0.8953	-0.2045	-0.5531	-0.5972	4.7445e-4
0.4000	0.5197	-1.4729	-1.3523	-0.3472	-1.0539	-1.2771	1.1384e-3
0.5000	0.6107	-1.9728	-2.0426	-0.5893	-2.0082	-2.7308	2.7314e-3
0.6000	0.7176	-2.6425	-3.0851	-1.0003	-3.8266	-5.8394	6.5535e-3
0.7000	0.8433	-3.5395	-4.6599	-1.6979	-7.2916	-12.4864	1.5724e-2

Effect of Choice of $\{\lambda_k\}$.

Consider a two-dimensional SRBM with covariance matrix $\sigma(x) = I$, drift vector $b(x) = [-1, 0]^T$ and reflection matrix

$$R = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}.$$

This example was considered in [25]. We consider the first moment of the x_1-coordinate. The exact value for this moment is known to be 0.5. We consider $\lambda_k = k^{-\alpha}$ and examine the influence of the choice of α on the numerical performance. The results are given in Figure 6.2.

Note that in Figure 6.2, we show that how $\int x_1\nu_n(dx)$ changes with n. This in some sense maybe misleading, since for different α, the actual time span and the computation time are quite different. For example, the actual time span and computing time for $\alpha = 0.9$ is 40.6 and 737s respectively, while the time for $\alpha = 0.1$ is $1.1347 * 10^5$ and 29400s. To make a better comparison, we summarize the value of $\int x_1\nu_n(dx)$ for different α with the same computing time 1 hour, and the results are shown in Table 6.2. In this example, the true value is 0.5, and $\alpha = 0.7$ gives the closest result.

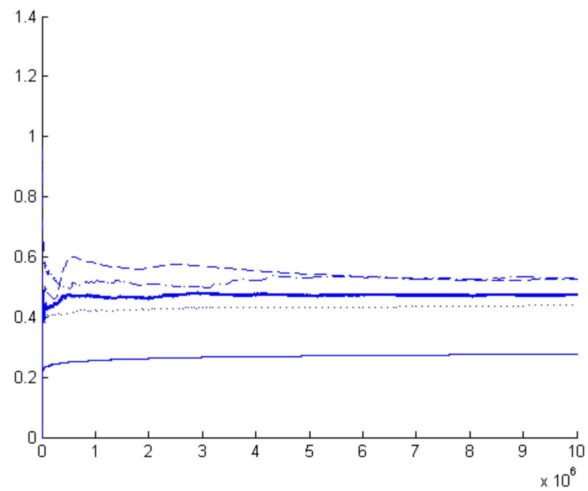

Figure 6.2: We consider time step sequence $\lambda_n = n^{-\alpha}$ with different choice of α and study the influence of α on numerical convergence. The thin solid line, the dotted line, the thick solid line, the dash-dot line, and the dashed line correspond to α =0.1, 0.3, 0.5, 0.7, and 0.9 respectively. The x-axis shows the value of n while the y-axis corresponds to $\int x_1 \nu_n(dx)$.

Table 6.2: Estimates for $\int x_1 \nu_n(dx)$ with same computing time of 1 hour for different α.

α	0.1	0.3	0.5	0.7	0.9
$\int x_1 \nu_n(dx)$	0.2565	0.4217	0.4681	0.4841	0.6218

An 8-d symmetric SRBM.

A SRBM is said to be symmetric if its covariance matrix Γ, drift vector μ and reflection matrix R are symmetric in the following sense: $\Gamma_{ij} = \Gamma_{ji} = \rho$ for $1 \le i < j \le d$, $\mu_i = -1$ for $1 \le i \le d$ and $R_{ij} = R_{ji} = -r$ for $1 \le i < j \le d$, where $r \ge 0$. The positiveness of Γ implies $-1/(d-1) < \rho < 1$ and the completely-$\mathcal{S}$ condition of R implies $r(d-1) < 1$. In this case, It is known (see [25]) that, the first moment of each of the component is the same, and is given by the following formula

$$m_1 = \frac{1 - (d-2)r + (d-1)r\rho}{2(1+r)}.$$

Here we take $d = 8$. Then the conditions on the data yield $-1/7 < \rho < 1$ and $0 \leq r < 1/7$. Letting ρ range through $\{-0.1, -0.05, 0, 0.2, 0.9\}$, and r take value 0.1, we obtain estimates of m_1 using algorithm in this work. We take $\lambda_k = k^{-\alpha}$, $\alpha = 0.5$ and $n = 10^7$. The results are shown in Table 6.3. The results show that as the correlation coefficient ρ approaches 1, the performance of the algorithm deteriorates. This may due to the fact that as ρ getting closer to 1, the covariance matrix becomes degenerate. The stationary distribution will be concentrated in a small space, which will be harder for the simulated trajectory to hit.

Table 6.3: Estimates for m_1 when $d = 8$.

ρ	-0.1	-0.05	0	0.2	0.9
Estimated Val.	0.131	0.137	0.163	0.414	3.205
True Val.	0.150	0.166	0.182	0.246	0.468

6.5 Appendix.

Lemma 6.5.1. *Let U be a random variable with bounded support. Suppose that* $\mathbf{E}U = 0$. *Then there exists $\alpha \in (0, \infty)$, such that*

$$\mathbf{E}e^{\lambda U} \leq e^{\alpha\lambda^2} \text{ for all } \lambda \in \mathbb{R}.$$

Proof. Without lots of generality we assume that $|U| \leq 1$.

Using the convexity of the function $e^{\lambda x}$, we have

$$e^{\lambda U} \leq \frac{U+1}{2}e^{\lambda} + \frac{1-U}{2}e^{-\lambda}.$$

Taking expectations in the above inequality and using Taylor's expansion, we have

$$\mathbf{E}e^{\lambda U} \leq \frac{e^{\lambda} + e^{-\lambda}}{2} \leq e^{\frac{\lambda^2}{2}}.$$

The lemma then follows on taking $\alpha = \frac{1}{2}$. $\square$

Bibliography

[1] R. Atar, A. Budhiraja, and P. Dupuis. On positive recurrence of constrained diffusion processes. *Annals of Probability*, pages 979–1000, 2001.

[2] R. Azencott. Grandes deviations et applications. *Ecole d'été de Probabilités de Saint-Flour VIII-1978*, pages 1–176, 1980.

[3] P. Baldi. Large deviations for diffusion processes with homogenization and applications. *The Annals of Probability*, 19(2):509–524, 1991.

[4] G. Basak, I. Hu, and C. Wei. Weak convergence of recursions. *Stochastic Processes and their Applications*, 68(1):65–82, 1997.

[5] H. Bessaih and A. Millet. Large deviation principle and inviscid shell models. *Electronic Journal of Probability*, 14:89, 2009.

[6] P. Billingsley. *Convergence of probability measures*, volume 493. Wiley-Interscience, 2009.

[7] A. Biswas and A. Budhiraja. Exit time and invariant measure asymptotics for small noise constrained diffusions. *Stochastic Processes and their Applications*, 121(5):899–924, 2011.

[8] V. Borkar and A. Budhiraja. Ergodic control for constrained diffusions: Characterization using HJB equations. *SIAM Journal on Control and Optimization*, 43(4):1467–1492, 2004.

[9] M. Boué and P. Dupuis. A variational representation for certain functionals of Brownian motion. *The Annals of Probability*, 26(4):1641–1659, 1998.

[10] A. Budhiraja. An ergodic control problem for constrained diffusion processes: existence of optimal Markov control. *SIAM Journal on Control and Optimization*, 42(2):532–558, 2003.

[11] A. Budhiraja, J. Chen, and P. Dupuis. Large deviations for stochastic partial differential equations driven by a Poisson random measure. *Stochastic Processes and their Applications*, 123(2):523–560, 2012.

[12] A. Budhiraja, J. Chen, and S. Rubenthaler. A numerical scheme for invariant distributions of constrained diffusions. *arXiv preprint arXiv:1205.5083*, 2012.

[13] A. Budhiraja and P. Dupuis. Simple necessary and sufficient conditions for the stability of constrained processes. *SIAM Journal on Applied Mathematics*, 59(5):1686–1700, 1999.

[14] A. Budhiraja and P. Dupuis. A variational representation for positive functionals of infinite dimensional Brownian motion. *Probability and Mathematical Statistics-Wroclaw University*, 20(1):39–61, 2000.

[15] A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *The Annals of Probability*, 36(4):1390–1420, 2008.

[16] A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for infinite dimensional stochastic dynamical systems. *The Annals of Probability*, 36(4):1390–1420, 2008.

[17] A. Budhiraja, P. Dupuis, and V. Maroulas. Large deviations for stochastic flows of diffeomorphisms. *Bernoulli*, 16(1):234–257, 2010.

[18] A. Budhiraja, P. Dupuis, and V. Maroulas. Variational representations for continuous time processes. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 47, pages 725–747. Institut Henri Poincaré, 2011.

[19] A. Budhiraja and C. Lee. Long time asymptotics for constrained diffusions in polyhedral domains. *Stochastic Processes and their Applications*, 117(8):1014–1036, 2007.

[20] A. Budhiraja and C. Lee. Stationary distribution convergence for generalized Jackson networks in heavy traffic. *Mathematics of Operations Research*, 34(1):45–56, 2009.

[21] A. Budhiraja and X. Liu. Stability of constrained Markov modulated diffusions, 2010.

[22] S. Cerrai and M. Freidlin. Averaging principle for a class of stochastic reaction–diffusion equations. *Probability Theory and Related Fields*, 144(1):137–177, 2009.

[23] I. Chueshov and A. Millet. Stochastic 2D hydrodynamical type systems: Well posedness and large deviations. *Applied Mathematics & Optimization*, 61(3):379–420, 2010.

[24] R. Cottle, J. Pang, and R. Stone. *The linear complementarity problem*. Number 60. Society for Industrial Mathematics, 2009.

[25] J. Dai and J. Harrison. Reflected Brownian motion in an orthant: numerical methods for steady-state analysis. *The Annals of Applied Probability*, pages 65–86, 1992.

[26] J. Dai and T. Kurtz. Characterization of the stationary distribution for a semimartingale reflecting Brownian motion in a convex polyhedron. *preprint*, 1994.

[27] J. Dai and T. Kurtz. A multiclass station with Markovian feedback in heavy traffic. *Mathematics of Operations Research*, 20(3):721–742, 1995.

[28] J. Dai and R. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions in convex polyhedrons. *Theory of Probability & Its Applications*, 40(1):1–40, 1995.

[29] A. de Acosta. Large deviations for vector-valued Lévy processes. *Stochastic Processes and their Applications*, 51(1):75–115, 1994.

[30] A. Du, J. Duan, and H. Gao. Small probability events for two-layer geophysical flows under uncertainty. *arXiv preprint arXiv:0810.2818*, 2008.

[31] J. Duan and A. Millet. Large deviations for the Boussinesq equations under random influences. *Stochastic Processes and their Applications*, 119(6):2052–2081, 2009.

[32] P. Dupuis. Large deviations analysis of some recursive algorithms with state dependent noise. *The Annals of Probability*, pages 1509–1536, 1988.

[33] P. Dupuis and R. Ellis. *A weak convergence approach to the theory of large deviations*, volume 902. Wiley-Interscience, 2011.

[34] P. Dupuis and H. Ishii. On Lipschitz continuity of the solution mapping to the Skorokhod problem, with applications. *Stochastics: An International Journal of Probability and Stochastic Processes*, 35(1):31–62, 1991.

[35] P. Dupuis and K. Ramanan. Convex duality and the Skorokhod problem. I. *Probability Theory and Related Fields*, 115(2):153–195, 1999.

[36] P. Dupuis and K. Spiliopoulos. Large deviations for multiscale diffusion via weak convergence methods. *Stochastic Processes and their Applications*, 122(4):1947–1987, 2011.

[37] P. Dupuis and H. Wang. Subsolutions of an Isaacs equation and efficient schemes

for importance sampling. *Mathematics of Operations Research*, 32(3):723–757, 2007.

[38] R. Durrett. *Probability: theory and examples*. Cambridge University Press, 2010.

[39] M. Freidlin. Fluctuations in dynamical systems with averaging (Russian) Dokl. *Akad. Nauk SSSR*, 226:273–276, 1976.

[40] M. Freidlin. The averaging principle and theorems on large deviations. *Russian Mathematical Surveys*, 33(5):117–176, 1978.

[41] M. Freidlin and R. Sowers. A comparison of homogenization and large deviations, with applications to wavefront propagation. *Stochastic Processes and their Applications*, 82(1):23–52, 1999.

[42] M. Freidlin and A. Wentzell. *Random perturbations of dynamical systems*, volume 260. Springer Verlag, 1998.

[43] T. Futagami, M. Yatsuzuka, and N. Tamai. FEM coupled with LP for water pollution control. *Journal of the Hydraulics Division*, 102(7):881–897, 1976.

[44] D. Gamarnik and A. Zeevi. Validity of heavy traffic steady-state approximations in generalized Jackson networks. *The Annals of Applied Probability*, 16(1):56–90, 2006.

[45] P. Hall and C. Heyde. *Martingale limit theory and its application.* Academic press New York, 1980.

[46] J. Harrison and M. Reiman. Reflected Brownian motion on an orthant. *The Annals of Probability*, 9(2):302–308, 1981.

[47] J. Harrison and R. Williams. Brownian models of open queueing networks with homogeneous customer populations. *Stochastics: An International Journal of Probability and Stochastic Processes*, 22(2):77–115, 1987.

[48] J. Harrison and R. Williams. Multidimensional reflected Brownian motions having exponential stationary distributions. *The Annals of Probability*, pages 115–137, 1987.

[49] N. Ikeda and S. Watanabe. Stochastic differential equations and diffusion processes, 1981.

[50] J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer-Verlag Berlin, 1987.

[51] G. Kallianpur, X. J., G. Hardy, and S. Ramasubramanian. The existence and uniqueness of solutions of nuclear space-valued stochastic differential equations driven by Poisson random measures. *Stochastics: An International Journal of Probability and Stochastic Processes*, 50(1-2):85–122, 1994.

[52] G. Kallianpur and J. Xiong. *Stochastic differential equations in infinite dimensional spaces*, volume 26. Inst of Mathematical Statistic, 1995.

[53] T. Kurtz. *Approximation of population processes*, volume 36. Society for Industrial Mathematics, 1981.

[54] T. Kurtz. A control formulation for constrained Markov processes. *Mathematics of Random Media*, 27:139–150, 1991.

[55] H. Kushner. *Heavy traffic analysis of controlled queueing and communication networks*, volume 47. Springer, 2001.

[56] H. Kushner. Large deviations for two-time-scale diffusions, with delays. *Applied Mathematics & Optimization*, 62(3):295–322, 2010.

[57] D. Lamberton and G. Pagès. Recursive computation of the invariant distribution of a diffusion. *Bernoulli*, 8(3):367–405, 2002.

[58] R. Liptser. Large deviations for two scaled diffusions. *Probability Theory and Related Fields*, 106(1):71–104, 1996.

[59] R. Liptser and A. Pukhalskii. Limit theorems on large deviations for semimartingales. *Stochastics: An International Journal of Probability and Stochastic Processes*, 38(4):201–249, 1992.

[60] R. Liptser, V. Spokoiny, and A. Veretennikov. Freidlin-Wentzell type large deviations for smooth processes. *Markov Process. Related Fields*, 8(4):611–636, 2002.

[61] W. Liu. Large deviations for stochastic evolution equations with small multiplicative noise. *Applied Mathematics & Optimization*, 61(1):27–56, 2010.

[62] A. Mandelbaum and G. Pats. State-dependent stochastic networks. Part I. Approximations and applications with continuous diffusion limits. *The Annals of Applied Probability*, 8(2):569–646, 1998.

[63] U. Manna, S. Sritharan, and P. Sundar. Large deviations for the stochastic shell model of turbulence. *NoDEA: Nonlinear Differential Equations and Applications*, 16(4):493–521, 2009.

[64] M. Pelletier. Weak convergence rates for stochastic approximation with application to multiple targets and simulated annealing. *Annals of Applied Probability*, pages 10–44, 1998.

[65] W. Peterson. A heavy traffic limit theorem for networks of queues with multiple customer types. *Mathematics of Operations Research*, 16(1):90–118, 1991.

[66] M. Reiman. Open queueing networks in heavy traffic. *Mathematics of Operations Research*, 9(3):441–458, 1984.

[67] M. Reiman and R. Williams. A boundary property of semimartingale reflecting Brownian motions. *Probability Theory and Related Fields*, 77(1):87–97, 1988.

[68] J. Ren and X. Zhang. Freidlin–Wentzell's large deviations for homeomorphism flows of non-Lipschitz SDEs. *Bulletin des Sciences Mathematiques*, 129(8):643–655, 2005.

[69] J. Ren and X. Zhang. Schilder theorem for the Brownian motion on the diffeomorphism group of the circle. *Journal of Functional Analysis*, 224(1):107–133, 2005.

[70] M. Röckner and T. Zhang. Stochastic evolution equations of jump type: existence, uniqueness and large deviation principles. *Potential Analysis*, 26(3):255–279, 2007.

[71] M. Röckner, T. Zhang, and X. Zhang. Large deviations for stochastic tamed 3D Navier-Stokes equations. *Applied Mathematics & Optimization*, 61(2):267–285, 2010.

[72] H. Royden and P. Fitzpatrick. *Real analysis*, volume 3. Prentice Hall Englewood Cliffs, NJ:, 1988.

[73] A. Skorokhod. *Asymptotic methods in the theory of stochastic differential equations*, volume 78. Amer Mathematical Society, 2008.

[74] S. Sritharan and P. Sundar. Large deviations for the two-dimensional Navier–Stokes equations with multiplicative noise. *Stochastic Processes and their Applications*, 116(11):1636–1659, 2006.

[75] D. Talay. Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics: An International Journal of Probability and Stochastic Processes*, 29(1):13–36, 1990.

[76] D. Talay. Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, 8(2):163–198, 2002.

[77] L. Taylor and R. Williams. Existence and uniqueness of semimartingale reflecting Brownian motions in an orthant. *Probability Theory and Related Fields*, 96(3):283–317, 1993.

[78] A. Veretennikov. Large deviations in averaging principle for stochastic differential equation systems (noncompact case). *Stochastics: An International Journal of Probability and Stochastic Processes*, 48(1-2):83–96, 1994.

[79] A. Veretennikov. On large deviations in the averaging principle for SDEs with a "full dependence". *Annals of Probability*, pages 284–296, 1999.

[80] A. Veretennikov. On large deviations for SDEs with small diffusion and averaging. *Stochastic Processes and their Applications*, 89(1):69–79, 2000.

[81] W. Wang and J. Duan. Reductions and deviations for stochastic partial differential equations under fast dynamical boundary conditions. *Stochastic Analysis and Applications*, 27(3):431–459, 2009.

[82] A. Weiss. *Invariant measures of diffusions in bounded domains*. PhD thesis, Ph. D. dissertation, New York University, 1981.

[83] A. Wentzell. *Limit theorems on large deviations for Markov stochastic processes*, volume 38. Springer, 1990.

[84] R. Williams. Diffusion approximations for open multiclass queueing networks: sufficient conditions involving state space collapse. *Queueing Systems*, 30(1):27–88, 1998.

[85] K. Yamada. Diffusion approximation for open state-dependent queueing networks in the heavy traffic situation. *The Annals of Applied Probability*, 5(4):958–982, 1995.

[86] D. Yang and Z. Hou. Large deviations for the stochastic derivative Ginzburg–Landau equation with multiplicative noise. *Physica D: Nonlinear Phenomena*, 237(1):82–91, 2008.

[87] G. Yin and Q. Zhang. *Continuous-time Markov chains and applications*. Springer, 1998.

[88] G. Yin and C. Zhu. *Hybrid switching diffusions: properties and applications*, volume 63. Springer, 2009.

[89] X. Zhang. Euler schemes and large deviations for stochastic Volterra equations with singular kernels. *Journal of Differential Equations*, 244(9):2226–2250, 2008.

[90] X. Zhang. Stochastic Volterra equations in Banach spaces and stochastic partial differential equation. *Journal of Functional Analysis*, 258(4):1361–1425, 2010.